

**SEARCH REQUEST FORM**

Scientific and Technical Information Center

Requester's Full Name: \_\_\_\_\_ Examiner #: \_\_\_\_\_ Date: \_\_\_\_\_  
 Art Unit: \_\_\_\_\_ Phone Number 30 \_\_\_\_\_ Serial Number: \_\_\_\_\_  
 Mail Box and Bldg/Room Location: \_\_\_\_\_ Results Format Preferred (circle): PAPER DISK E-MAIL

**If more than one search is submitted, please prioritize searches in order of need.**

\*\*\*\*\*

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc. if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: \_\_\_\_\_

Inventors (please provide full names): \_\_\_\_\_

Earliest Priority Filing Date: \_\_\_\_\_

*\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

Jan Delavai  
 Reference Librarian  
 Biotechnology & Chemical Librarian  
 CM1 1E07 - 703-308-4498  
 jan.delavai@uspto.gov

\*\*\*\*\*  
**STAFF USE ONLY**

	Type of Search	Vendors and cost where applicable
Searcher: <u>Jan</u>	NA Sequence (#) _____	STN _____
Searcher Phone #: <u>4498</u>	AA Sequence (#) <input checked="" type="checkbox"/>	Dialog _____
Searcher Location: _____	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: <u>12/4/02</u>	Bibliographic _____	Dr.Link _____
Date Completed: <u>12/4/02</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep & Review Time: _____	Fulltext _____	Sequence Systems <input checked="" type="checkbox"/>
Clerical Prep Time: <u>15</u>	Patent Family _____	WWW/Internet _____
Online Time: <u>+ 20</u>	Other _____	Other (specify) _____



# STIC SEARCH RESULTS FEEDBACK FORM

## Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher* or contact:

Mary Hale, Information Branch Supervisor  
Remsen Bldg. 01 D86  
571-272-2507

## Voluntary Results Feedback Form

➤ I am an examiner in Workgroup:  Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s).
- ☐ Non-Patent Literature  
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC-Biotech-Chem Library, Remsen Bldg.



GenCore version 5.1.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 13:05:05 ; Search time 15 Seconds  
(without alignments)  
198.678 Million cell updates/sec

Title: US-09-362-731A-1

Perfect score: 165

Sequence: 1 QYIKANSKFIGITELGGHEIKKVLVPGCHGS 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 6612

Minimum DB seq length: 0  
Maximum DB seq length: 31

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	35	21.2	26	2 B46451	RNA-directed RNA p
2	32	19.4	31	2 S53153	gene X protein - h
3	31	18.8	21	2 S69371	duodenase - bovine
4	30	18.2	21	2 A27719	trypsin (EC 3.4.21
5	30	18.2	24	2 B48368	N5,N10-methenyltet
6	30	18.2	25	2 I51151	alpha actin - qual
7	30	18.2	26	2 A28108	atrial gland pepti
8	29	17.6	20	2 S19618	globin - polychaet
9	29	17.6	23	2 G81890	hypothetical prote
10	29	17.6	24	2 A44008	lethal peptide II
11	29	17.6	31	2 A70219	hypothetical prote
12	28.5	17.3	20	2 PQ0732	superoxide dismuta
13	28.5	17.3	24	2 A47209	histone H-1-MDBP-2
14	28.5	17.3	27	2 C44908	chitinase (EC 3.2.
15	28	17.0	20	2 S57286	translation elonga
16	28	17.0	21	2 J00386	nitrile hydratase
17	28	17.0	26	2 A33413	cytochrome f - tur
18	28	17.0	31	2 S23181	formaldehyde dehyd
19	28	17.0	31	2 S13663	cellulase (EC 3.2.
20	28	17.0	31	2 S53233	gene X protein - h
21	27	16.4	7	2 S42407	gramicidin S synth
22	27	16.4	13	2 S74130	NADH oxidase - Gia
23	27	16.4	21	2 S68023	nucleoside-diphos
24	27	16.4	24	2 B44008	lethal peptide I -
25	27	16.4	30	2 S63531	hypothetical prote
26	27	16.4	31	2 S53190	gene X protein - h
27	26.5	16.1	20	2 P50188	superoxide dismuta
28	26.5	16.1	26	2 S16181	calpain (EC 3.4.22
29	26	15.8	14	2 I39753	nitrogenase (EC 1.

30 26 15.8 17 2 A61334  
31 26 15.8 21 2 C39543  
32 26 15.8 23 2 S51188  
33 26 15.8 23 2 B30988  
34 26 15.8 26 2 C44379  
35 26 15.8 27 2 S52355  
36 26 15.8 30 2 E31461  
37 26 15.8 31 2 J70381  
38 26 15.8 31 2 S53208  
39 26 15.8 31 2 E82856  
40 25.5 15.5 31 2 S53162  
41 25 15.2 22 2 S18130  
42 25 15.2 23 2 S60569  
43 25 15.2 24 2 S11384  
44 25 15.2 25 2 PT0088  
45 25 15.2 27 2 PC2337

#### ALIGNMENTS

##### RESULT 1

B46451  
RNA-directed RNA polymerase (EC 2.7.7.48) - parainfluenza virus type 3 (strain 47885)  
C:Species: parainfluenza virus type 3  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 18-Jun-1999  
C:Accession: B46451  
R:Storey, D.G.; Cote, M.J.; Dimock, K.; Kang, C.Y.  
Intervirology 27, 69-80, 1987  
A:Title: Nucleotide sequence of the coding and flanking regions of the human parainfl  
A:Reference number: A46451; MUID:88032139; PMID:2822598  
A:Accession: B46451  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-26 <S70>  
A:Cross-references: GB:M20402; NID:q332720; PIDN:AAA46857.1; PID:q332722  
C:Superfamily: parainfluenza virus RNA-directed RNA polymerase  
C:Keywords: ATP; nucleotidyltransferase

Query Match 21.2%; Score 35; DB 2; Length 26;  
Best Local Similarity 35.3%; Pred. No. 2.3e+02;  
Matches 6; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 13 TELGGHEIKKVLVPGCH 29  
|| : |||  
Db 3 TESNNGTVSDILYPECH 19

##### RESULT 2

S53153  
gene X protein - hepatitis B virus (isolate patient Usai'89) (fragment)  
C:Species: hepatitis B virus, HBV  
A:Variety: isolate patient Usai'89  
C:Date: 08-Jul-1995 #sequence\_revision 03-Aug-1995 #text\_change 26-Aug-1999  
C:Accession: S53153  
R:Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.  
submitted to the EMBL Data Library, March 1995  
A:Reference number: S53112  
A:Accession: S53153  
A:Molecule type: DNA  
A:Residues: 1-31 <LAI>  
A:Cross-references: EMBL:X85270; NID:g736037; PIDN:CAA59557.1; PID:g736038  
A:Experimental source: isolate patient Usai'89  
C:Genetics:  
A:Gene: X  
C:Superfamily: hepatitis B virus gene X protein

Query Match 19.4%; Score 32; DB 2; Length 31;  
Best Local Similarity 42.3%; Pred. No. 7.4e+02;  
Matches 11; Conservative 3; Mismatches 10; Indels 2; Gaps 2;

QY 5 ANSKFIGITELGGHEIKKVLVPG-CH 29  
|:|:| : | || | | | |

Db 2 ADSRLI-IFVLGGRHKLVCAPACH 26

RESULT 3  
S69371  
duodenase - bovine (fragment)  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 21-Nov-1998  
C:Accession: S69371  
R:Zamolodchikova, T.S.; Vorotyntseva, T.I.; Antonov, V.K.  
Eur. J. Biochem. 227, 866-872, 1995  
A:Title: Duodenase, a new serine protease of unusual specificity from bovine duodenal mu  
A:Reference number: S69371; MUID:95172075; PMID:7867648  
A:Accession: S69371  
A:Molecule type: protein  
A:Residues: 1-21 <ZAM>  
C:Superfamily: trypsin; trypsin homology

Query Match 18.8%; Score 31; DB 2; Length 21;  
Best Local Similarity 71.4%; Pred. No. 6.9e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 15 LGGHEIK 21  
:|||||  
DB 2 IGGHEAK 8

RESULT 4  
A27719  
trypsin (EC 3.4.21.4) precursor - marbled lungfish (fragment)  
C:Species: Protopterus aethiopicus (marbled lungfish)  
C:Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 08-Nov-1996  
C:Accession: A27719  
R:Hermanson, M.A.; Tye, R.W.; Reeck, G.R.; Neurath, H.; Walsh, K.A.  
FEBS Lett. 14, 222-224, 1971  
A:Title: Comparison of the amino terminal sequences of bovine, dogfish, and lungfish try  
A:Reference number: A27719  
A:Accession: A27719  
A:Molecule type: protein  
A:Residues: 1-21 <HER>  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; protein digestion; serine proteinase

Query Match 18.2%; Score 30; DB 2; Length 21;  
Best Local Similarity 50.0%; Pred. No. 9.6e+02;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 15 LGGHEIKKVLVP 26  
:|||||  
DB 9 VGGYECPRKHXP 20

RESULT 5  
B48368  
N5,N10-methylnetetrahydromethanopterin cyclohydrolase - Methanosarcina barkeri (fragment  
C:Species: Methanosarcina barkeri  
C:Date: 19-Nov-1993 #sequence\_revision 18-Nov-1994 #text\_change 18-Nov-1994  
C:Accession: B48368  
R:Klein, A.R.; Breitung, J.; Linder, D.; Stetter, K.O.; Thauer, R.K.  
Arch. Microbiol. 159, 213-219, 1993  
A:Title: N5,N10-methylnetetrahydromethanopterin cyclohydrolase from the extremely thermo  
rom the extremely thermophilic Methanopyrus kandleri.  
A:Reference number: A48368; MUID:93243882; PMID:8481088  
A:Contents: Fusaro D  
A:Accession: B48368  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-24 <KLE>  
A>Note: sequence extracted from NCBI backbone (NCBIP:130470)

Query Match 18.2%; Score 30; DB 2; Length 24;  
Best Local Similarity 33.3%; Pred. No. 1.1e+03;  
Matches 5; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVL 24  
:|||||  
DB 2 ISVNEGSGNVIEEML 16

## RESULT 6

I51151

alpha actin - quail (fragment)

C:Species: Coturnix coturnix (quail)

C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 13-Aug-1999

C:Accession: I51151

R:Hastings, K.E.M.; Emerson, C.P.

Proc. Natl. Acad. Sci. U.S.A. 79, 1553-1557, 1982

A:Title: cDNA clone analysis of six co-regulated mRNAs encoding skeletal muscle contr

A:Reference number: I51151; MUID:82174638; PMID:6951196

A:Accession: I51151

A&gt;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-25 &lt;HAS&gt;

A:Cross-references: GB:J00957; NID:g213596; PIDN:AAA49491.1; PID:g213597

C:Superfamily: actin

Query Match 18.2%; Score 30; DB 2; Length 25;  
Best Local Similarity 54.5%; Pred. No. 1.2e+03;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 9 FIGITELGGHE 19

:|||||

DB 11 FIGMESAGIHE 21

## RESULT 7

A28108

atrial gland peptide D1 - California sea hare (fragment)

C:Species: Aplysia californica (California sea hare)

C:Date: 28-Aug-1989 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999

C:Accession: A28108

R:Nagle, G.T.; Painter, S.D.; Blankenship, J.E.; Kurosky, A.

J. Biol. Chem. 263, 9223-9237, 1988

A:Title: Proteolytic processing of egg-laying hormone-related precursors in Aplysia.

A:Reference number: A92727; MUID:88243802; PMID:3379066

A:Accession: A28108

A:Molecule type: protein

A:Residues: 1-26 &lt;NAG&gt;

Query Match 18.2%; Score 30; DB 2; Length 26;  
Best Local Similarity 50.0%; Pred. No. 1.2e+03;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 15 LGGHEIKKVL 24

:|||||

DB 6 LGGHQLTEEL 15

## RESULT 8

S19618

globin - polychaete (Eudistyllia vancouveri) (fragment)

N:Alternate names: chlorocruorin

C:Species: Eudistyllia vancouveri

C:Date: 19-Mar-1997 #sequence\_revision 10-Oct-1997 #text\_change 05-Dec-1998

C:Accession: S19618

R:Qabar, A.N.; Stern, M.S.; Walz, D.A.; Chiu, J.T.; Timkovich, R.; Wall, J.S.; Kapp,

J. Mol. Biol. 222, 1109-1129, 1991

A:Title: Hierarchy of globin complexes. The quaternary structure of the extracellular

A:Reference number: S19532; MUID:92106333; PMID:1762147

A:Accession: S19618

A:Molecule type: protein

A:Residues: 1-20 &lt;QAB&gt;

A:Experimental source: plume

C:Complex: dodecamers, each consisting of a trimer of tetramers of globin chains; dode

C:Keywords: calcium; dodecamer; heme; homotetramer; oxygen carrier



Query Match 17.6%; Score 29; DB 2; Length 20;  
 Best Local Similarity 35.3%; Pred. No. 1.3e+03;  
 Matches 6; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 14 ELGGHEIKKVLVPGCHG 30  
 || |::: | |  
 DB 1 ELSSVEVKRIDANGKKG 17

RESULT 9  
 G81890  
 hypothetical protein NMA1224 [imported] - Neisseria meningitidis (strain Z2491 serogroup  
 C:Species: Neisseria meningitidis  
 C:Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 02-Feb-2001  
 C:Accession: G81890  
 R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel  
 ; Holroyd, S.; Jagers, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,  
 Nature 404, 502-506, 2000  
 A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.  
 A:Reference number: A81775; MUID:20222556; PMID:10761919  
 A:Accession: G81890  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-23 <PAR>  
 A:Cross-references: GB:AL162755; GB:AL157959; NID:g7379742; PIDN:CAB84484.1; PID:g737991  
 A:Experimental source: serogroup A, strain Z2491  
 C:Genetics:  
 A:Gene: NMA1224

Query Match 17.6%; Score 29; DB 2; Length 23;  
 Best Local Similarity 75.0%; Pred. No. 1.5e+03;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 17 GHEIKKVL 24  
 |||| | |  
 DB 13 GHEIAKKL 20

RESULT 10  
 A44008  
 lethal peptide II - Wagler's pit viper  
 N:Alternate names: waglerin SL-II  
 C:Species: Trimeresurus wagleri (Wagler's pit viper)  
 C:Date: 27-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 01-Mar-1996  
 C:Accession: A44008; B39559  
 R:Schmidt, J.J.; Weinstein, S.A.; Smith, L.A.  
 Toxicon 30, 1027-1036, 1992  
 A:Title: Molecular properties and structure-function relationships of lethal peptides fr  
 A:Reference number: A44008; MUID:93069257; PMID:1440639  
 A:Accession: A44008  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-24 <SCH>  
 A:Experimental source: venom  
 A>Note: sequence extracted from NCBI backbone (NCBIP:119455)  
 R:Weinstein, S.A.; Schmidt, J.J.; Bernheimer, A.W.; Smith, L.A.  
 Toxicon 29, 227-236, 1991  
 A:Title: Characterization and amino acid sequences of two lethal peptides isolated from  
 A:Reference number: A39559; MUID:91263101; PMID:2048140  
 A:Accession: B39559  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 3-24 <WEI>

Query Match 17.6%; Score 29; DB 2; Length 24;  
 Best Local Similarity 40.0%; Pred. No. 1.5e+03;  
 Matches 6; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 15 LGGHEIKKVLVPGCH 29  
 ||| : | | |  
 DB 2 LGGKPDLPFCYPPCH 16

RESULT 11  
 A70219  
 hypothetical protein BBB21 - Lyme disease spirochete plasmid B/cp26  
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
 C:Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 21-Jul-2000  
 C:Accession: A70219  
 R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh  
 son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu  
 ; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.  
 Nature 390, 580-586, 1997  
 A:Authors: Smith, H.O.; Venter, J.C.  
 A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.  
 A:Reference number: A70100; MUID:98065943; PMID:9403685  
 A:Accession: A70219  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-31 <KLE>  
 A:Cross-references: GB:AE000792; NID:g3253098; PIDN:AAC66340.1; PID:g2689909; TIGR:BB  
 A:Experimental source: strain B31  
 C:Genetics:  
 A:Genome: plasmid

Query Match 17.6%; Score 29; DB 2; Length 31;  
 Best Local Similarity 54.5%; Pred. No. 2e+03;  
 Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 QYTKANSKFTIG 11  
 | | | | | |  
 DB 19 QIFKDNSSFFG 29

RESULT 12  
 PQ0732  
 superoxide dismutase (EC 1.15.1.1) (Cu-Zn) 5.7/16K [similarity] - rice (fragment)  
 C:Species: Oryza sativa (rice)  
 C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
 C:Accession: PQ0732  
 R:Komatsu, S.; Kajiwara, H.; Hirano, H.  
 Theor. Appl. Genet. 86, 935-942, 1993  
 A:Title: A rice protein library: a data-file of rice proteins separated by two-dimens  
 A:Reference number: PQ0696  
 A:Accession: PQ0732  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-20 <KOM>  
 C:Keywords: metalloprotein; oxidoreductase

Query Match 17.3%; Score 28.5; DB 2; Length 20;  
 Best Local Similarity 28.6%; Pred. No. 1.5e+03;  
 Matches 6; Conservative 7; Mismatches 7; Indels 1; Gaps 1;

QY 5 ANSKFIGITELGGHEIKKVLV 25  
 | | : : | | : : | |  
 DB 1 ATKKAVAVL-IGTHQVEGVV 20

RESULT 13  
 A47209  
 histone H-1-MDBP-2 - chicken (fragments)  
 C:Species: Gallus gallus (chicken)  
 C:Date: 27-Jun-1994 #sequence\_revision 27-Jun-1994 #text\_change 27-Jun-1994  
 C:Accession: A47209  
 R:Jost, J.P.; Hofsteenge, J.  
 Proc. Natl. Acad. Sci. U.S.A. 89, 9499-9503, 1992  
 A:Title: The repressor MDBP-2 is a member of the histone H1 family that binds prefere  
 A:Reference number: A47209; MUID:93028489; PMID:1409659  
 A:Accession: A47209  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-24 <JOS>

Query Match 17.3%; Score 28.5; DB 2; Length 24;  
 Best Local Similarity 33.3%; Pred. No. 1.8e+03;

Matches 6; Conservative 5; Mismatches 0; Indels 7; Gaps 1;

QY 12 ITEL-----GGHEIKK 22  
:|:|:  
Db 7 VTELITKALAAGGYDVEK 24  
|:|:|:|

## RESULT 14

C44908  
chitinase (EC 3.2.1.14), 47K - Streptomyces olivaceoviridis (fragment)  
C:Species: Streptomyces olivaceoviridis  
C:Date: 01-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 16-Feb-1997  
C:Accession: C44908  
R:Romaguera, A.; Menge, U.; Breves, R.; Diekmann, H.  
J. Bacteriol. 174, 3450-3454, 1992  
A:Title: Chitinases of Streptomyces olivaceoviridis and significance of processing for  
A:Reference number: A44908; PMID:92276319; PMID:1592803  
A:Accession: C44908  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-27 <ROM>  
A:Experimental source: ATCC 11238  
A:Note: sequence extracted from NCBI backbone (NCBIP:104591)  
C:Superfamily: Streptomyces chitinase chi40  
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 17.3%; Score 28.5; DB 2; Length 27;

Best Local Similarity 44.4%; Pred. No. 2.1e+03;

Matches 8; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

QY 5 ANSKFEG-ITELGGHEIK 21  
|:|:|:|:|:|:  
Db 1 AGSKVGVYFTWGYDVK 18  
|:|:|:|:|:|:

## RESULT 15

S57286  
translation elongation factor aEF-1 beta - Sulfolobus solfataricus (fragments)  
C:Species: Sulfolobus solfataricus  
C:Date: 27-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 17-Mar-1999  
C:Accession: S57286  
R:Arcari, P.; Raimo, G.; Iannicciello, G.; Gallo, M.; Bocchini, V.  
Biochim. Biophys. Acta 1263, 86-88, 1995  
A:Title: The first nucleotide sequence of an archaeal elongation factor 1-beta gene.  
A:Reference number: S57286; PMID:95359209; PMID:7632739  
A:Accession: S57286  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-20 <ARC>

Query Match 17.0%; Score 28; DB 2; Length 20;

Best Local Similarity 71.4%; Pred. No. 1.8e+03;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 10 IGITELG 16  
:|:|:|:  
Db 13 VGITRLG 19  
|:|:|:|:|:|:

Search completed: December 4, 2002, 13:07:24

Job time : 16 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 12:55:15 ; Search time 11 Seconds

(without alignments)  
116.888 Million cell updates/sec

Title: US-09-362-731A-1

Perfect score: 165

Sequence: 1 QYIKANSKFIGITELGGHEIKKVLVPGCHGS 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 2000

Minimum DB seq length: 0

Maximum DB seq length: 31

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	30	18.2	21	1 TRYP_PROAT	P35051 prototenus
2	28.5	17.3	27	1 CHI2_STROI	P25116 streptomyc
3	28	17.0	21	1 NBLA_ACISP	P33036 acinetobact
4	28	17.0	27	1 CXO7_CONCN	P58916 conus conso
5	28	17.0	28	1 GUN_SCHCO	P81190 schizophyll
6	28	17.0	31	1 GUN2_SCLSC	P21834 sclerotinia
7	27	16.4	21	1 NDK_CANAL	Q9ur66 candida alb
8	26	15.8	14	1 TKNL_SCHGR	P82470 schistocerc
9	26	15.8	20	1 FRB3_LITIN	P56249 litoria inf
10	26	15.8	22	1 PSP2_PHYPA	P80661 physcomitre
11	26	15.8	25	1 CXOB_CONCT	P58918 conus catus
12	26	15.8	31	1 PETL_OENHO	Q9mtk4 oenothera h
13	25.5	15.5	26	1 MGN_CHICK	P50594 gallus gall
14	25	15.2	14	1 SAP2_ARBPB	P11760 arabacia pun
15	25	15.2	21	1 MDH_RHOPA	P80458 rhodopsudo
16	25	15.2	22	1 CR33_LITCE	P56240 litoria cae
17	25	15.2	22	1 CR34_LITCE	P56241 litoria cae
18	25	15.2	22	1 LP2_TRIWA	P58930 trimeresuru
19	25	15.2	23	1 TYPH_LACCA	P19663 lactobacill
20	25	15.2	25	1 CYC4_VIBFI	P80891 vibrio fisc
21	25	15.2	26	1 CXOC_CONCT	P58919 conus catus
22	25	15.2	28	1 LPL_ECOLI	P81440 aspergillus
23	25	15.2	28	1 PHYB_ASPTI	P81440 aspergillus
24	25	15.2	30	1 UC35_MAIZE	P80641 zea mays (m
25	25	15.2	31	1 ETFD_PARDE	P55932 paracoccus
26	24.5	14.8	19	1 CHI0_CLOPA	P81338 clostridium
27	24	14.5	11	1 TKC2_CALVO	P57104 rana tempor
28	24	14.5	13	1 TMLL_RANTE	P57104 rana tempor
29	24	14.5	15	1 PH2_PERAM	P82695 periplaneta
30	24	14.5	17	1 PH3_PERAM	P82696 periplaneta
31	24	14.5	25	1 AU51_LITRA	P82401 litoria ran
32	24	14.5	26	1 RL20_BUCAP	P46246 buchnera ap
33	24	14.5	29	1 CXOC_CONMA	P37300 conus magus

34	24	14.5	29	1	CXOD_CONMA	Q26350 conus magus
35	24	14.5	29	1	PTL_ENTFA	P23530 enterococcu
36	24	14.5	30	1	ALAT_CHIVI	P38026 chinchilla
37	24	14.5	30	1	DMS3_PHYSA	P80279 phyllomedus
38	24	14.5	30	1	FIBR_PANIN	P22775 panulirus i
39	24	14.5	30	1	PCCA_MYXXA	P81185 myxococcus
40	24	14.5	30	1	PG20_AKBKL	P82143 agkistrodon
41	24	14.5	31	1	RL21_STRTR	P48956 streptococc
42	23.5	14.2	20	1	DNAK_CLOPA	P81341 clostridium
43	23	13.9	8	1	COW2_CONPU	P58785 conus purpu
44	23	13.9	15	1	CX1B_CONBE	P58624 conus betul
45	23	13.9	17	1	TRP2_LEUMA	P81733 leucophaea

## ALIGNMENTS

RESULT 1  
TRYP\_PROAT  
ID TRYP\_PROAT STANDARD; PRT; 21 AA.  
AC P35051;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Trypsin precursor (EC 3.4.21.4) (Fragment).  
OS Prototenus aethiopicus (Marbled lungfish).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Dipnoi; Lepidosireniformes; Protopteridae; Protopterus.  
OX NCBI\_TaxID=7886;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Pancreas;  
RA Hermodson M.A., Tye R.W., Reeck G.R., Neurath H., Walsh K.A.;  
RT "Comparison of the amino terminal sequences of bovine, dogfish, and  
lungfish trypsinogens.";  
RL FEBS Lett. 14:222-224(1971).  
CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Arg-I-Xaa, Lys-I-Xaa.  
CC -|- SUBCELLULAR LOCATION: Extracellular.  
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
DR PIR; A27719; A27719.  
DR MEROPS; S01.151; -.  
DR InterPro; IPR001254; Ser\_protease\_Try.  
DR PROSITE; PS50240; TRYPSIN\_DOM; PARTIAL.  
DR PROSITE; PS00134; TRYPSIN\_HIS; PARTIAL.  
DR PROSITE; PS00135; TRYPSIN\_SER; PARTIAL.  
KW Hydrolase; Serine protease; zymogen.  
FT PROPEP 1 7 ACTIVATION PEPTIDE.  
FT CHAIN 8 >21 TRYPSIN.  
FT NON\_TER 21 21  
SQ SEQUENCE 21 AA; 2454 MW; 88E2FB1D130729D CRC64;

Query Match 18.2%; Score 30; DB 1; Length 21;  
Best Local Similarity 50.0%; Pred. No. 3.4e+02;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 15 LGGHEIKKVLVLP 26  
:||:| | | |  
Db 9 VGGYECGPKHXP 20

RESULT 2  
CHI2\_STROI  
ID CHI2\_STROI STANDARD; PRT; 27 AA.  
AC P29116;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 01-DEC-1992 (Rel. 24, Last annotation update)  
DE Chitinase 47 kDa (EC 3.2.1.14) (Fragment).  
OS Streptomyces olivaceoviridis (Streptomyces corchorusii).  
OC Bacteria; Actinobacteria; Actinobacteriales; Streptomyces.  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1921;  
RN [1]

```

RP SEQUENCE.
RC STRAIN=ATCC 11238;
RX MEDLINE=92276319; PubMed=1592803;
RA Romaguera A., Menge U., Breves R., Diekmann H.;
RT "Chitinases of Streptomyces olivaceoviridis and significance of
RT processing for multiplicity.";
RL J. Bacteriol. 174:3450-3454(1992).
CC -|- FUNCTION: ABLE TO CLEAVE CHITIN OLIGOMERS FROM N=3 TO 6.
CC -|- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
CC acetyl-D-glucosamine polymers of chitin.
CC -|- SUBUNIT: HOMODIMER.
CC -|- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC HYDROLASES).
DR PIR; C44908;
DR InterPro; IPR001579; Chitinase_18/2.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE_18; PARTIAL.
KW Hydrolase; Glycosidase; Chitin degradation.
FT NON_TER 27 27
SQ SEQUENCE 27 AA; 3172 MW; E9549D5E273DFEE9 CRC64;

Query Match 17.0%; Score 28.5; DB 1; Length 27;
Best Local Similarity 44.4%; Pred. No. 7.4e+02;
Matches 8; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

QY 5 ANSKFEG-ITELGGHEIK 21
   | | | : | | | : | |
DB 1 AGSRVGVFTWGYDRK 18

RESULT 3
NRLA_ACISP
ID NRLA_ACISP STANDARD; PRT; 21 AA.
AC P3036;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Nitrilase (EC 3.5.5.1) (Fragment).
OS Acinetobacter sp. (strain AK226).
OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
OC Acinetobacter.
OX NCBI_TaxID=472;
RN [1]
RP SEQUENCE.
RX MEDLINE=91345837; PubMed=1369128;
RA Yamamoto K., Komatsu K.;
RT "Purification and characterization of nitrilase responsible for the
RT enantioselective hydrolysis from Acinetobacter sp. AK 226.";
RL Agric. Biol. Chem. 55:1459-1466(1991).
CC -|- FUNCTION: ACTS ON MANY KINDS OF NITRILE COMPOUNDS SUCH AS
CC ALIPHATIC, AROMATIC, AND HETEROCYCLIC MONONITRILES OR DINITRILES.
CC PREFERS S-(-)-2-(4'-ISOBUTYLPHENYL)-PROPIONITRILE TO R-(+)-2-(4'-
CC ISOBUTYLPHENYL)-PROPIONITRILE AS THE SUBSTRATE.
CC -|- CATALYTIC ACTIVITY: A nitrile + H(2)O = a carboxylate + NH(3).
CC -|- SIMILARITY: BELONGS TO THE NITRILASE FAMILY.
DR PIR; J00386;
DR InterPro; IPR000132; N/CN_hydrolase.
DR PROSITE; PS00920; NITRIL_CHT_1; PARTIAL.
DR PROSITE; PS00921; NITRIL_CHT_2; PARTIAL.
KW Hydrolase.
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2223 MW; 5FA741C41EAC619B CRC64;

Query Match 17.0%; Score 28; DB 1; Length 21;
Best Local Similarity 45.5%; Pred. No. 6.8e+02;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 IKANSKFIGIT 13
   : | | | | : |
DB 1 VSYNSKFLAAT 11

```

```

RESULT 4
CX07_CONCN STANDARD; PRT; 27 AA.
AC P5916;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Omega-conotoxin CnVIIA.
OS Conus consors (Singed cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=101297;
RN [1]
RP SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.
RC TISSUE=Venom;
RX PubMed=11724570;
RA Favreau P., Gilles N., Lamthanh H., Bournaud R., Shimahara T.,
RA Bouet F., Laboute P., Letourneux Y., Menez A., Molgo J., Le Gall F.;
RT "A new omega-conotoxin that targets N-type voltage-sensitive calcium
RT channels with unusual specificity.";
RL Biochemistry 40:14567-14575(2001).
CC -|- FUNCTION: Omega-conotoxins act at presynaptic membranes, they bind
CC and block voltage-sensitive calcium channels (VSCC). This toxin
CC blocks N-type calcium channels, but unexpectedly, does not show
CC any blocking activity at amphibian neuromuscular junction. Causes
CC shaking activity, and, at higher doses, causes mild tremors when
CC injected intracerebroventricularly into mice. Causes paralysis,
CC and, at higher doses, causes death when injected intramuscularly
CC into fish.
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -|- MASS SPECTROMETRY: MW=2847.74; METHOD=Electrospray.
CC -|- SIMILARITY: BELONGS TO THE O-SUPERFAMILY OF CONOTOXINS. OMEGA-TYPE
CC FAMILY.
KW Presynaptic neurotoxin; Neurotoxin; Toxin; Calcium channel inhibitor;
KW Hydroxylation; Amidation.
FT BINDING 13 13 ESSENTIAL FOR CALCIUM CHANNEL BINDING (BY
FT SIMILARITY).
FT DISULFID 1 16 BY SIMILARITY.
FT DISULFID 8 20 BY SIMILARITY.
FT DISULFID 15 27 BY SIMILARITY.
FT MOD_RES 7 7 HYDROXYLATION.
FT MOD_RES 27 27 AMIDATION.
SQ SEQUENCE 27 AA; 2839 MW; B9DEFD149iF2CB4A CRC64;

Query Match 17.0%; Score 28; DB 1; Length 27;
Best Local Similarity 33.3%; Pred. No. 8.7e+02;
Matches 5; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 17 GHEIKVLVPGCHGS 31
   | : : : | | | |
DB 5 GAPCTRLMYDCCHGS 19

RESULT 5
GUN_SCHCO
ID GUN_SCHCO STANDARD; PRT; 28 AA.
AC P81190;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Endoglucanase (EC 3.2.1.4) (Endo-1,4-beta-glucanase) (Cellulase)
DE (Fragment).
OS Schizophyllum commune (Bracket fungus).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Agaricales; Schizophyllaceae; Schizophyllum.
OX NCBI_TaxID=5334;
RN [1]
RP SEQUENCE.
RX MEDLINE=97459758; PubMed=9315718;
RA Clarke A.J., Drummel-Smith J., Yaguchi M.;
RT "Identification of the catalytic nucleophile in the cellulase from
RT Schizophyllum commune and assignment of the enzyme to family 5,

```

OS	Candida albicans (Yeast).
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC	Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX	NCBI_TaxID=5476;
[1]	
RN	SEQUENCE, AND CHARACTERIZATION.
RX	MEDLINE=96019963; PubMed=7487065;
RA	Blondi R.M., Veron M., Walz K., Passeron S.;
RA	"Candida albicans nucleoside-diphosphate kinase: purification and
RT	characterization.";
RL	Arch. Biochem. Biophys. 323:187-194(1995).
CC	-i- FUNCTION: MAJOR ROLE IN THE SYNTHESIS OF NUCLEOSIDE TRIPHOSPHATES
CC	OTHER THAN ATP (BY SIMILARITY).
CC	-i- CATALYTIC ACTIVITY: ATP + nucleoside diphosphate = ADP +
CC	nucleoside triphosphate.
CC	-i- SUBUNIT: HOMOHEXAMER.
CC	-i- SUBCELLULAR LOCATION: Cytoplasmic.
CC	-i- SIMILARITY: BELONGS TO THE NDK FAMILY.
DR	HSP; P22887; ILWX.
DR	InterPro; IPR001564; NDK.
DR	PFam; PF00334; NDK; 1.
DR	PROSITE; PS00459; NDP_KINASES; 1.
KW	Transferase; Kinase; ATP-binding.
FT	NON_TER 1 1
FT	ACT_SITE 4 4
FT	NON_TER 21 21
FT	BY SIMILARITY.
FT	SEQUENCE 21 AA; 2379 MW; 9DABB3A325947001 CRC64;
QY	Query Match 16.4%; Score 27; DB 1; Length 21;
DB	Best Local Similarity 100.0%; Pred. No. 9.5e+02;
DB	Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps
QY	28 CHGS 31
DB	
DB	3 CHGS 6
RESULT 8	
TKNL_SCHGR	
ID_TKNL_SCHGR	STANDARD; PRT; 14 AA.
AC	PC2470;
DT	30-MAY-2000 (Rel. 39, Created)
DT	30-MAY-2000 (Rel. 39, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)
DE	Tachykinin-1 (Scg-midgut-TK).
OS	Schistocerca gregaria (Desert locust).
OC	Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC	Insecta; Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caellifera;
OC	Acridomorpha; Acridoidea; Acrididae; Cyrtacanthacridinae;
OC	Schistocerca.
OX	NCBI_TaxID=7010;
[1]	
RN	SEQUENCE, AND MASS SPECTROMETRY.
RP	TISSUE=Midgut;
RC	MEDLINE=20050081; PubMed=105811195;
RX	Veelaert D., Baggerman G., Derua R., Waelkens E., Meeusen T.,
RA	Vande Water G., De Loof A., Schoofs L.;
RA	"Identification of a new tachykinin from the midgut of the desert
RT	locust, Schistocerca gregaria, by ESI-Qq-oe-TOF mass spectrometry.";
RL	Biochem. Biophys. Res. Commun. 266:237-242(1999).
CC	-i- FUNCTION: MYOACTIVE PEPTIDE. INCREASES THE AMPLITUDE AND FREQUENCY
CC	OF SPONTANEOUS CONTRACTIONS AND TONUS OF HINDGUT MUSCLE.
CC	-i- TISSUE SPECIFICITY: MIDGUT.
CC	-i- MASS SPECTROMETRY: MW=1493.79; METHOD=Electrospray.
CC	-i- SIMILARITY: SIMILAR TO THE COCKROACH LENTRP 3, A TACHYKININ-
CC	RELATED PEPTIDE ALSO CONFINED TO THE MIDGUT.
KW	Tachykinin; Neuropeptide; Amidation.
FT	MOD_RES 14 14
FT	AMIDATION
FT	SEQUENCE 14 AA; 1496 MW; CA4C578C0169FC72 CRC64;
QY	Query Match 15.8%; Score 26; DB 1; Length 14;
DB	Best Local Similarity 50.0%; Pred. No. 8.9e+02;
DB	Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps

```

QY 22 KVLVPGCHGS 31
| : | : | :
Db 4 KRAVPGFYGT 13

RESULT 9
PRE3_LITIN
ID PRE3_LITIN STANDARD; PRT; 20 AA.
AC P56249;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Prenatin 3.
OS Litoria infrarenata (Giant tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC Pelodyadinae; Litoria.
OX NCBI_TaxID=61195;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;
RX MEDLINE=97368637; PubMed=9225251;
RA Raftery M.J., Waugh R.J., Bowie J.H., Wallace J.C., Tyler M.J.;
RT "The structures of the frenatin peptides from the skin secretion of
RL J. Pept. Sci. 2:117-124(1996).
CC -!- FUNCTION: WIDE SPECTRUM ANTIMICROBIAL PEPTIDE.
CC -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN PAROTOID AND/OR ROSTRAL
CC GLANDS.
KW Antibiotic; Amphibian skin.
SQ SEQUENCE 20 AA; 1956 MW; 7E4ABE30EAL7B20C CRC64;

Query Match 15.8%; Score 26; DB 1; Length 20;
Best Local Similarity 35.7%; Pred. No. 1.3e+03;
Matches 5; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 11 GITELGGHGHIKKVL 24
| : | : | : |
Db 1 GLMSVLGHVGNVL 14

RESULT 10
PSP2_PHYPA
ID PSP2_PHYPA STANDARD; PRT; 22 AA.
AC P80661;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Oxygen-evolving enhancer protein 2 (OEE2) (22 kDa subunit of oxygen
DE evolving system of photosystem II) (Fragments).
OS Physcomitrella patens (Moss).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
OC Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
OX NCBI_TaxID=3218;
RN [1]
RP SEQUENCE.
RC TISSUE=Protonema;
RX MEDLINE=97275459; PubMed=9129336;
RA Kasten B., Buck F., Nuske J., Reski R.;
RT "Cytokinin affects nuclear- and plastome-encoded energy-converting
RT plastid enzymes.";
RL Planta 201:261-272(1997).
CC -!- FUNCTION: MAY BE INVOLVED IN THE REGULATION OF PHOTOSYSTEM II.
CC -!- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane; associated
CC with the photosystem II complex.
CC -!- INDUCTION: BY LIGHT.
CC -!- SIMILARITY: BELONGS TO THE PSPB FAMILY.
KW Photosynthesis; Photosystem II; Chloroplast; Thylakoid; Membrane;
FT Multigene family.
FT NON_CONS 15 16
FT NON_TER 22 22
SQ SEQUENCE 22 AA; 2316 MW; E1553CC575C54BB1 CRC64;

```

```

Query Match 15.8%; Score 26; DB 1; Length 22;
Best Local Similarity 50.0%; Pred. No. 1.4e+03;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 5 ANSKFICITELG 16
| : | : | : |
Db 10 AGNGFTXITEYG 21

RESULT 11
CXOB_CONCT
ID CXOB_CONCT STANDARD; PRT; 25 AA.
AC P58918;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Omega-conotoxin CVIB.
OS Conus catus (Cat cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=101291;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RC TISSUE=Venom;
RX PubMed=10938268;
RA Lewis R.J., Nielsen K.J., Craik D.J., Loughnan M.L., Adams D.A.,
RA Sharpe I.A., Luchian T., Adams D.J., Bond T., Thomas L., Jones A.,
RA Matheson J.-L., Drinkwater R., Andrews P.R., Alewood P.F.;
RT "Novel omega-conotoxins from Conus catus discriminate among neuronal
RT calcium channel subtypes.";
RL J. Biol. Chem. 275:35335-35344(2000).
CC -!- FUNCTION: Omega-conotoxins act at presynaptic membranes, they bind
CC and block voltage-sensitive calcium channels (VSCC) (By
CC similarity). This toxin blocks N-, P-, and Q-type calcium
CC channels.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -!- SIMILARITY: BELONGS TO THE O-SUPERFAMILY OF CONOTOXINS. OMEGA-TYPE
CC FAMILY.
KW Presynaptic neurotoxin; Neurotoxin; Toxin; Calcium channel inhibitor;
KW Amidation.
FT DISULFID 1 16 BY SIMILARITY.
FT DISULFID 8 20 BY SIMILARITY.
FT DISULFID 15 25 BY SIMILARITY.
FT MOD_RES 25 25 AMIDATION.
SQ SEQUENCE 25 AA; 2717 MW; D4IA9E5FSAFA9552 CRC64;

Query Match 15.8%; Score 26; DB 1; Length 25;
Best Local Similarity 33.3%; Pred. No. 1.6e+03;
Matches 5; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 17 GHEIKKVLVPGCHGS 31
| : | : | : |
Db 5 GASCRKTMVDCRCGS 19

RESULT 12
PETL_OENHO
ID PETL_OENHO STANDARD; PRT; 31 AA.
AC Q9MTK4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cytochrome B6-F complex subunit VI (Cytochrome b6f complex subunit
DE petL).
GN PETL.
OS Oenothera hookeri (Hooker's evening primrose).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Myrtales; Onagraceae; Oenothera.
OX NCBI_TaxID=85636;

```

```

RN  SEQUENCE FROM N.A.
RP  SUPRAIN-cv. Johansen.
RC  MEDLINE=20309318; PubMed=10852478;
RA  Hupfer H., Swiatek M., Hornung S., Herrmann R.G., Maier R.M.,
RX  Chiu W.-L., Sears B.;
RT  "Complete nucleotide sequence of the Oenothera elata plastid
RT  chromosome, representing plastome I of the five distinguishable
RT  Euoenothera plastomes.";
RL  Mol. Gen. Genet. 263:581-585(2000).
CC  -!- FUNCTION: THE CYTOCHROME B6-F COMPLEX FUNCTIONS IN THE LINEAR
CC  CROSS-MEMBRANE TRANSPORT OF ELECTRONS BETWEEN PHOTOSYSTEM II AND
CC  I, AS WELL AS IN CYCLIC ELECTRON FLOW AROUND PHOTOSYSTEM I. PETL
CC  IS IMPORTANT FOR PHOTOAUTOTROPHIC GROWTH AS WELL AS FOR ELECTRON
CC  TRANSFER EFFICIENCY AND STABILITY OF THE CYTOCHROME B6-F COMPLEX
CC  (By similarity).
CC  -!- SUBCELLULAR LOCATION: Thylakoid membrane-associated (By
CC  similarity).
CC  -!- SIMILARITY: BELONGS TO THE PETL FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AJ271079; CAB67175.1;
KW  Electron transport; Chloroplast; Respiratory chain; Transmembrane;
KW  Thylakoid.
FT  TRANSMEM 4 24 POTENTIAL.
SQ  SEQUENCE 31 AA; 3415 MW; A015C65D2F325493 CRC64;

Query Match 15.8%; Score 26; DB 1; Length 31;
Best Local Similarity 57.18; Pred. No. 2e+03;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY  9 FIGITEL 15
DB  22 FIGLTKI 28

RESULT 13
MGN_CHICK
ID  MGN_CHICK STANDARD; PRT; 26 AA.
AC  P50594;
DT  01-OCT-1996 (Rel. 34, Created)
DT  01-OCT-1996 (Rel. 34, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Mago nashi protein homolog (Fragment).
GN  MAGOH.
OS  Gallus gallus (Chicken).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC  Gallus.
OX  NCBI_TaxID=9031;
RN  NCBI_TaxID=9031;
RP  SEQUENCE FROM N.A.
RC  TISSUE=Heart;
RA  Shi N., Antin P., Flink I.L., Morkin E.;
RL  Submitted (MAY-1994) to the EMBL/GenBank/DBSJ databases.
CC  -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC  -!- SIMILARITY: BELONGS TO THE MAGO NASHI FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----

```

```

DR  EMBL; L29089; -: NOT_ANNOTATED_CDS.
DR  InterPro: IPR004023; Mago_nashi.
DR  Pfam: PF02792; Mago_nashi; 1.
KW  Nuclear protein.
FT  NON_TER 26
SQ  SEQUENCE 26 AA; 3062 MW; AF82470DE9E29230 CRC64;

Query Match 15.5%; Score 25.5; DB 1; Length 26;
Best Local Similarity 25.0%; Pred. No. 2e+03;
Matches 6; Conservative 4; Mismatches 7; Indels 7; Gaps 1;

QY  2 YKANSKFTIGITELGGHEIKKVLV 25
DB  10 YVGHKCKF-----GHELNVLCI 26

RESULT 14
SAP2_ARBP
ID  SAP2_ARBP STANDARD; PRT; 14 AA.
AC  P11760;
DT  01-OCT-1989 (Rel. 12, Created)
DT  01-OCT-1989 (Rel. 12, Last sequence update)
DT  01-MAR-1992 (Rel. 21, Last annotation update)
DE  Resact (Sperm-activating peptide) (SAP-IIA).
OS  Arbacia punctulata (Punctate sea urchin).
OC  Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC  Echinoidea; Euechinoidea; Echinacea; Arbacoidea; Arbaciidae; Arbacia.
OX  NCBI_TaxID=7641;
RN  [1]
RP  SEQUENCE.
RC  TISSUE=Egg;
RX  MEDLINE=85054981; PubMed=6150045;
RA  Suzuki N., Shimomura H., Radany E.W., Ramarao C.S., Ward G.E.,
RA  Bentley J.K., Garbers D.L.;
RT  "A peptide associated with eggs causes a mobility shift in a major
RT  plasma membrane protein of spermatozoa.";
RL  J. Biol. Chem. 259:14874-14879(1984).
RN  [2]
RP  DISULFIDE BOND.
RX  MEDLINE=92097763; PubMed=1756838;
RA  Yoshino K.-I., Takao T., Shimonishi Y., Suzuki N.;
RT  "Determination of the amino acid sequence of an intramolecular
RT  disulfide linkage-containing sperm-activating peptide by tandem mass
RT  spectrometry.";
RL  FEBS Lett. 294:179-182(1991).
CC  -!- FUNCTION: CAUSE STIMULATION OF SPERM RESPIRATION AND MOTILITY
CC  THROUGH INTRACELLULAR ALKALINIZATION, TRANSIENT ELEVATIONS OF
CC  CAMP, CGMP AND CALCIUM LEVELS IN SPERM CELLS, AND TRANSIENT
CC  ACTIVATION AND SUBSEQUENT INACTIVATION OF THE MEMBRANE FORM OF
CC  GUANYLATE CYCLASE.
CC  -!- SIMILARITY: SMALL TO S.PURPURATUS SPERACT.
KW  Amidation.
FT  DISULFID 1 8 AMIDATION.
FT  MOD_RES 14 14
SQ  SEQUENCE 14 AA; 1246 MW; 39745AA33EBE41B8 CRC64;

Query Match 15.2%; Score 25; DB 1; Length 14;
Best Local Similarity 80.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY  26 PGCHG 30
DB  6 PGCVG 10

RESULT 15
MDH_RHOPA
ID  MDH_RHOPA STANDARD; PRT; 21 AA.
AC  P80458;
DT  15-JUL-1998 (Rel. 36, Created)
DT  15-JUL-1998 (Rel. 36, Last sequence update)
DT  30-MAY-2000 (Rel. 39, Last annotation update)
DE  Malate dehydrogenase (EC 1.1.1.37) (Fragment).

```

GN MDH  
OS Rhodopseudomonas palustris.  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
CC Bradyrhizobium group; Rhodopseudomonas.  
OX NCBI\_TaxID=1076;  
RN [1]  
RP SEQUENCE.  
RA Naterstad K., Synstad B., Sirevag R.;  
RL Submitted (SEP-1996) to the SWISS-PROT data bank.  
CC -!- CATALYTIC ACTIVITY: (S)-malate + NAD(+) = oxaloacetate + NADH.  
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE LDH FAMILY. MDH SUBFAMILY.  
DR InterPro; IPR001252; MDH.  
DR PROSITE; PS00068; MDH; PARTIAL.  
KW Oxidoreductase; Tricarboxylic acid cycle; NAD.  
FT NON\_TER 21 21  
SQ SEQUENCE 21 AA; 2053 MW; FA97098FF856205F CRC64;

Query Match 15.2%; Score 25; DB 1; Length 21;  
Best Local Similarity 62.5%; Pred. No. 1.9e+03;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 10 IGITELGG 17  
|||  
Db 5 IGSGEIGG 12

Search completed: December 4, 2002, 13:06:28  
Job time : 13 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 13:04:30 ; Search time 28 Seconds

(without alignments)  
228.124 Million cell updates/sec

Title: US-09-362-731A-1

Perfect score: 165

Sequence: 1 QYIKANSKFIGITELGGHEIKKVLVPGCHGS 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 15780

Minimum DB seq length: 0

Maximum DB seq length: 31

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPTREMBL\_21.\*  
2: sp-archaea.\*  
3: sp-bacteria.\*  
4: sp-fungi.\*  
5: sp-human.\*  
6: sp-invertebrate.\*  
7: sp-mammal.\*  
8: sp-mhc.\*  
9: sp-organelle.\*  
10: sp-phage.\*  
11: sp-plant.\*  
12: sp-rodent.\*  
13: sp-virus.\*  
14: sp-vertebrate.\*  
15: sp-unclassified.\*  
16: sp-rvirus.\*  
17: sp-bacteriap.\*  
17: sp-archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36.5	22.1	30	Q9R4I6	Q9r4i6 mycoplasma
2	36	21.8	31	Q9BE10	Q9bei10 monodelphis
3	36	21.8	31	Q95M83	Q95m83 equus caball
4	35.5	21.5	30	Q9R4I5	Q9r4i5 mycoplasma
5	33	20.0	22	Q9TRB1	Q9trb1 bos taurus
6	33	20.0	30	Q93GF6	Q93gf6 staphylococ
7	32	19.4	27	Q9QIE5	Q9qie5 hepatitis c
8	32	19.4	27	Q9QIE7	Q9qie7 hepatitis c
9	32	19.4	27	Q9QIE4	Q9qie4 hepatitis c
10	32	19.4	27	Q9QIE3	Q9qie3 hepatitis c
11	32	19.4	27	Q9QIE2	Q9qie2 hepatitis c
12	32	19.4	27	Q9QIE1	Q9qie1 hepatitis c
13	32	19.4	27	Q9QID7	Q9qid7 hepatitis c
14	32	19.4	27	Q9QID6	Q9qid6 hepatitis c
15	32	19.4	27	Q9QID4	Q9qid4 hepatitis c
16	32	19.4	27	Q9QID3	Q9qid3 hepatitis c

17	32	19.4	27	12	Q9QID1	Q9qid1 hepatitis c
18	32	19.4	27	12	Q9QID0	Q9qid0 hepatitis c
19	32	19.4	27	12	Q9QHK0	Q9qh0 hepatitis c
20	32	19.4	27	12	Q9QHJ9	Q9qh9 hepatitis c
21	32	19.4	27	12	Q9QHJ8	Q9qh8 hepatitis c
22	32	19.4	27	12	Q9QHJ6	Q9qh6 hepatitis c
23	32	19.4	27	12	Q9QHJ1	Q9qh1 hepatitis c
24	32	19.4	27	12	Q9QHJ0	Q9qh0 hepatitis c
25	32	19.4	27	12	Q9QH17	Q9qh17 hepatitis c
26	32	19.4	28	10	Q944P1	Q944p1 manihot esc
27	32	19.4	31	4	Q96C35	Q96c35 homo sapien
28	32	19.4	31	12	Q67974	Q67974 hepatitis b
29	31.5	19.1	27	6	Q29382	Q29382 sus scrofa
30	31	18.8	18	6	Q9TRQ2	Q9trq2 bos taurus
31	31	18.8	19	4	Q9UCC1	Q9ucc1 homo sapien
32	31	18.8	25	11	P70382	P70382 mus musculu
33	31	18.8	27	12	Q37180	Q37180 hepatitis c
34	31	18.8	27	12	Q9QHI8	Q9qh8 hepatitis c
35	31	18.8	31	2	Q93GF7	Q93gf7 staphylococ
36	30.5	18.5	29	2	Q9ACD9	Q9acd9 vibrio salm
37	30	18.2	21	2	O07840	O07840 rhodobacter
38	30	18.2	25	13	Q92191	Q92191 coturnix co
39	30	18.2	27	12	Q9QH83	Q9qh83 hepatitis c
40	30	18.2	27	12	Q9QH82	Q9qh82 hepatitis c
41	30	18.2	27	12	Q9QH81	Q9qh81 hepatitis c
42	30	18.2	27	12	Q9QH80	Q9qh80 hepatitis c
43	30	18.2	27	12	Q9QH79	Q9qh79 hepatitis c
44	30	18.2	27	12	Q9QH78	Q9qh78 hepatitis c
45	30	18.2	27	12	Q9QH77	Q9qh77 hepatitis c

## ALIGNMENTS

### RESULT 1

Q9R4I6 PRELIMINARY; PRT; 30 AA.  
ID Q9R4I6  
AC Q9R4I6;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE Arginine deiminase (EC 3.5.3.6) (Fragment).  
OS Mycoplasma arginini.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2094;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=96042668; PubMed=7591961;  
RA Takaku H., Matsumoto M., Misawa S., Miyazaki K.;  
RT "Anti-tumor activity of arginine deiminase from Mycoplasma argini and  
RT its growth-inhibitory mechanism.";  
RL Jpn. J. Cancer Res. 86:840-846(1995).  
SQ SEQUENCE 30 AA; 3389 MW; 66DA98752F6D6694 CRC64;

Query Match 22.1%; Score 36.5; DB 2; Length 30;  
Best Local Similarity 48.1%; Pred. No. 2.8e+02;  
Matches 13; Conservative 5; Mismatches 2; Indels 7; Gaps 3;

QY 6 NSKFIGI---TELGGHEIKKVLV--PG 27  
:||| || :||| || :||| ||

Db 4 DSKFKGIHVYSEIG--ELESVLVHEPG 28

### RESULT 2

ID Q9BE10 PRELIMINARY; PRT; 31 AA.  
AC Q9BE10;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Omega globin (Fragment).  
OS Monodelphis domestica (Short-tailed grey opossum).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Metatheria; Didelphimorphia; Didelphidae; Monodelphis.  
 OX NCBI\_TaxID=13616;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21107677; PubMed=11158601;  
 RA Wheeler D., Hope R., Cooper S.J., Dolman G., Webb G.C., Bottema C.D.,  
 RA Gooley A.A., Goodman M., Holland R.A.;  
 RT "An orphaned mammalian beta-globin gene of ancient evolutionary  
 origin";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:1101-1106(2001).  
 CC -!- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.  
 DR EMBL; AY014771; AAK11486.1; -;  
 DR HSSP; P02096; 1FDH.  
 DR InterPro; IPR000971; Globin.  
 DR Pfam; PF00042; globin; 1.  
 DR PROSITE; PS01033; GLOBIN; 1.  
 KW Heme; Oxygen transport.  
 FT NON\_TER 1 31  
 FT NON\_TER 31 31  
 SQ SEQUENCE 31 AA; 3547 MW; 0353D72F3681617D CRC64;  
 Query Match 21.8%; Score 36; DB 6; Length 31;  
 Best Local Similarity 37.5%; Pred. No. 3.4e+02;  
 Matches 6; Conservative 5; Mismatches 5; Indels 0; Gaps 0;  
 QY 10 IGITELGGHEIKKVLV 25  
 Db 10 VDIEKLGHTLSRLV 25  
 RESULT 3  
 Q95M83 PRELIMINARY; PRT; 31 AA.  
 AC Q95M83;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Pulmonary surfactant associated protein B (Fragment).  
 OS Equus caballus (Horse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
 OX NCBI\_TaxID=9796;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21314992; PubMed=11421942;  
 RA Shubitowski D.M., Venta P.J., Douglass C.L., Zhou R.-X., Ewart S.L.;  
 RT "Polymorphism identification within 50 equine gene-specific sequence  
 tagged sites";  
 RL Anim. Genet. 32:78-78(2001).  
 DR EMBL; AY008812; AAG40948.1; -;  
 DR InterPro; IPR000004; SapB.  
 DR ProDom; PD001732; SapB; 1.  
 FT NON\_TER 1 31  
 FT NON\_TER 31 31  
 SQ SEQUENCE 31 AA; 3671 MW; 17C43179D3685D17 CRC64;  
 Query Match 21.8%; Score 36; DB 6; Length 31;  
 Best Local Similarity 75.0%; Pred. No. 3.4e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 22 KVLVPGCH 29  
 Db 6 KLLVPGCH 13  
 RESULT 4  
 Q9R4I5 PRELIMINARY; PRT; 30 AA.  
 ID Q9R4I5;  
 AC Q9R4I5;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)

DE Arginine deiminase (EC 3.5.3.6) (Fragment).  
 OC Mycoplasma hominis.  
 OS Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
 OC Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2098;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=96042668; PubMed=7591961;  
 RA Takaku H., Matsumoto M., Misawa S., Miyazaki K.;  
 RT "Anti-tumor activity of arginine deiminase from Mycoplasma argini and  
 RT its growth-inhibitory mechanism";  
 RL Jpn. J. Cancer Res. 86:840-846(1995).  
 SQ SEQUENCE 30 AA; 3389 MW; 93DC43752F6D6022 CRC64;  
 Query Match 21.5%; Score 35.5; DB 2; Length 30;  
 Best Local Similarity 48.1%; Pred. No. 3.9e+02;  
 Matches 13; Conservative 5; Mismatches 2; Indels 7; Gaps 3;  
 QY 6 NSKFIGI---TELGGEIKKVLV--PG 27  
 Db 4 DSKFNGIHVYSEIG--ELETVLVHEPG 28  
 RESULT 5  
 Q9TRB1 PRELIMINARY; PRT; 22 AA.  
 ID Q9TRB1;  
 AC Q9TRB1;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE ATP-dependent 20 S proteasome activator (Fragment).  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=94342244; PubMed=8063704;  
 RA DeMartino G.N., Moomaw C.R., Zagnitko O.P., Proske R.J., Chu-Ping M.,  
 RA Afendis S.J., Swaffield J.C., Slaughter C.A.;  
 RT "PA700, an ATP-dependent activator of the 20 S proteasome, is an  
 RT ATPase containing multiple members of a nucleotide-binding protein  
 family";  
 RL J. Biol. Chem. 269:20878-20884(1994).  
 SQ SEQUENCE 22 AA; 2195 MW; FE139D0ACF4AE9BE CRC64;  
 Query Match 20.0%; Score 33; DB 6; Length 22;  
 Best Local Similarity 50.0%; Pred. No. 6.6e+02;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 QY 11 GITELGGHEI 20  
 Db 8 GILDAGGHNV 17  
 RESULT 6  
 Q93GF6 PRELIMINARY; PRT; 30 AA.  
 ID Q93GF6;  
 AC Q93GF6;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE AurB.  
 GN AURB.  
 OS Staphylococcus aureus.  
 OC Plasmid pRJ6.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
 OC Staphylococcus.  
 OX NCBI\_TaxID=1280;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21424233; PubMed=11531330;

RA Netz D.J.A., Sahl H.-G., Marcolino R., Nascimento J.S., Oliveira S.S.,  
RA Soares M.B., Bastos M.C.F.;  
RT "Molecular characterisation of aureocin A70, a multi-peptide  
bacteriocin isolated from *Staphylococcus aureus*.";  
RL J. Mol. Biol. 311:939-949(2001).  
RW EMBL; AF241888; AAK73553.1; -.  
KW Plasmid.  
SQ SEQUENCE 30 AA; 2797 MW; 4B60654093F92DFA CRC64;

Query Match 20.0%; Score 33; DB 2; Length 30;  
Best Local Similarity 53.8%; Pred. No. 9.3e+02;  
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 5 ANSKFIGITELGG 17

Db 3 AVAKFLGRLALGG 15

RESULT 7

Q9QIE5 ID Q9QIE5 PRELIMINARY; PRT; 27 AA.

AC Q9QIE5; 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Polypeptide (Fragment).  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]

RP SEQUENCE FROM N.A.  
RA Sandres K., Dubois M., Pasquier C., Izopet J.;  
RT "The genetic heterogeneity of hypervariable region 1 of the viral  
genome and the sensitivity of hepatitis C virus to interferon alpha  
therapy.";  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF166632; AAD52290.1; -.

FT NON\_TER 1  
FT NON\_TER 27  
SQ SEQUENCE 27 AA; 2695 MW; 84055F98C6EB5B7A CRC64;

Query Match 19.4%; Score 32; DB 12; Length 27;  
Best Local Similarity 46.2%; Pred. No. 1.2e+03;  
Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 7 SKFIGITELGGHE 19

Db 14 SAFAGFVRLGPHQ 26

RESULT 8

Q9QIE7 ID Q9QIE7 PRELIMINARY; PRT; 27 AA.

AC Q9QIE7; 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Genome polypeptide [Contains: envelope glycoprotein E2 (GP68) (GP70)  
(NSI)] (Fragment).  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]

RP SEQUENCE FROM N.A.  
RA Sandres K., Dubois M., Pasquier C., Izopet J.;  
RT "The genetic heterogeneity of hypervariable region 1 of the viral  
genome and the sensitivity of hepatitis C virus to interferon alpha  
therapy.";  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF166630; AAD52288.1; -.

KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;

KW Polyprotein; Transmembrane.  
FT NON\_TER 1  
FT NON\_TER 27  
SQ SEQUENCE 27 AA; 2727 MW; 8CA75E43B6EB5B61 CRC64;

Query Match 19.4%; Score 32; DB 12; Length 27;  
Best Local Similarity 46.2%; Pred. No. 1.2e+03;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 7 SKFIGITELGGHE 19

Db 14 SAFAGFVRLGPHQ 26

RESULT 9

Q9QIE4 ID Q9QIE4 PRELIMINARY; PRT; 27 AA.

AC Q9QIE4; 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Genome polypeptide [Contains: envelope glycoprotein E2 (GP68) (GP70)  
(NSI)] (Fragment).  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]

RP SEQUENCE FROM N.A.  
RA Sandres K., Dubois M., Pasquier C., Izopet J.;  
RT "The genetic heterogeneity of hypervariable region 1 of the viral  
genome and the sensitivity of hepatitis C virus to interferon alpha  
therapy.";  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF166633; AAD52291.1; -.

KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
KW Polyprotein; Transmembrane.  
FT NON\_TER 1  
FT NON\_TER 27  
SQ SEQUENCE 27 AA; 2727 MW; 8CA75E43B6EB5B61 CRC64;

Query Match 19.4%; Score 32; DB 12; Length 27;  
Best Local Similarity 46.2%; Pred. No. 1.2e+03;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 7 SKFIGITELGGHE 19

Db 14 SAFAGFVRLGPHQ 26

RESULT 10

Q9QIE3 ID Q9QIE3 PRELIMINARY; PRT; 27 AA.

AC Q9QIE3; 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Genome polypeptide [Contains: envelope glycoprotein E2 (GP68) (GP70)  
(NSI)] (Fragment).  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]

RP SEQUENCE FROM N.A.  
RA Sandres K., Dubois M., Pasquier C., Izopet J.;  
RT "The genetic heterogeneity of hypervariable region 1 of the viral  
genome and the sensitivity of hepatitis C virus to interferon alpha  
therapy.";  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF166634; AAD52292.1; -.

KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
KW Polyprotein; Transmembrane.

```
FT NON_TER 1 1
SQ SEQUENCE 27 AA: 2727 MW: 8CA75E43B6EB5B61 CRC64;

Query Match 19.4%; Score 32; DB 12; Length 27;
Best Local Similarity 46.2%; Pred. No. 1.2e+03;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 7 SKFIGITELGGHE 19
  I I I: I I I:
Db 14 SAFAGLRLGPHQ 26

RESULT 11
Q9QIE2 PRELIMINARY; PRT; 27 AA.
AC Q9QIE2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70)
DE (NS1)] (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sandres K., Dubois M., Pasquier C., Izopet J.;
RT "The genetic heterogeneity of hypervariable region 1 of the viral
RT genome and the sensitivity of hepatitis C virus to interferon alpha
RT therapy.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF166635; AAD52293.1; -.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 1 1
FT NON_TER 27 27
SQ SEQUENCE 27 AA: 2727 MW: 8CA75E43B6EB5B61 CRC64;

Query Match 19.4%; Score 32; DB 12; Length 27;
Best Local Similarity 46.2%; Pred. No. 1.2e+03;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 7 SKFIGITELGGHE 19
  I I I: I I I:
Db 14 SAFAGLRLGPHQ 26

RESULT 12
Q9QIE1 PRELIMINARY; PRT; 27 AA.
AC Q9QIE1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70)
DE (NS1)] (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sandres K., Dubois M., Pasquier C., Izopet J.;
RT "The genetic heterogeneity of hypervariable region 1 of the viral
RT genome and the sensitivity of hepatitis C virus to interferon alpha
RT therapy.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF166636; AAD52294.1; -.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 1 1
FT NON_TER 27 27
SQ SEQUENCE 27 AA: 2727 MW: 8CA75E43B6EB5B61 CRC64;

Query Match 19.4%; Score 32; DB 12; Length 27;
Best Local Similarity 46.2%; Pred. No. 1.2e+03;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 7 SKFIGITELGGHE 19
  I I I: I I I:
Db 14 SAFAGLRLGPHQ 26

RESULT 13
Q9QID7 PRELIMINARY; PRT; 27 AA.
AC Q9QID7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70)
DE (NS1)] (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sandres K., Dubois M., Pasquier C., Izopet J.;
RT "The genetic heterogeneity of hypervariable region 1 of the viral
RT genome and the sensitivity of hepatitis C virus to interferon alpha
RT therapy.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF166640; AAD52298.1; -.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 1 1
FT NON_TER 27 27
SQ SEQUENCE 27 AA: 2727 MW: 8CA75E43B6EB5B61 CRC64;

Query Match 19.4%; Score 32; DB 12; Length 27;
Best Local Similarity 46.2%; Pred. No. 1.2e+03;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 7 SKFIGITELGGHE 19
  I I I: I I I:
Db 14 SAFAGLRLGPHQ 26

RESULT 14
Q9QID6 PRELIMINARY; PRT; 27 AA.
AC Q9QID6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70)
DE (NS1)] (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sandres K., Dubois M., Pasquier C., Izopet J.;
RT "The genetic heterogeneity of hypervariable region 1 of the viral
RT genome and the sensitivity of hepatitis C virus to interferon alpha
RT therapy.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF166641; AAD52299.1; -.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 1 1
FT NON_TER 27 27
SQ SEQUENCE 27 AA: 2727 MW: 8CA75E43B6EB5B61 CRC64;
```

SQ SEQUENCE 27 AA; 2727 MW; 8CA75E43B6EB5B61 CRC64;  
Query Match 19.4%; Score 32; DB 12; Length 27;  
Best Local Similarity 46.2%; Pred. NO. 1.2e+03;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 7 SKFIGITELGGHE 19  
| | | : | | | :  
Db 14 SAFAGLRLGPHQ 26

## RESULT 15

Q9QIDA  
ID Q9QIDA PRELIMINARY; PRT; 27 AA.  
AC Q9QIDA;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70) (NS1)] (Fragment).  
DE (NS1)] (Fragment).  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Sandres K., Dubois M., Pasquier C., Izopet J.;  
RT "The genetic heterogeneity of hypervariable region 1 of the viral genome and the sensitivity of hepatitis C virus to interferon alpha therapy.";  
RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF166643; AAD52301.1; -.  
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
KW Polyprotein; Transmembrane.  
FT NON\_TER 1  
FT NON\_TER 27  
SQ SEQUENCE 27 AA; 2727 MW; 8CA75E43B6EB5B61 CRC64;

Query Match 19.4%; Score 32; DB 12; Length 27;  
Best Local Similarity 46.2%; Pred. NO. 1.2e+03;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 7 SKFIGITELGGHE 19  
| | | : | | | :  
Db 14 SAFAGLRLGPHQ 26

Search completed: December 4, 2002, 13:07:03  
Job time : 29 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 12:54:30 : Search time 34 Seconds

(without alignments)  
121.493 Million cell updates/sec

Title: US-09-362-731A-1

Perfect score: 165

Sequence: 1 QYIKANSKFCITELGGHEIKKVLPGCHGS 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 352784

Minimum DB seq length: 0

Maximum DB seq length: 31

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

```

1: /SID22/qcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/qcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/qcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/qcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/qcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/qcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/qcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/qcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/qcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/qcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/qcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/qcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/qcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/qcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/qcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/qcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/qcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/qcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/qcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/qcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/qcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/qcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/qcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	165	100.0	31	21 AAY82632	Tetanus toxoid T c
2	86	52.1	29	16 AAR35561	Igt CH4 region con
3	84	50.9	25	16 AAR77122	Dermatophagoides p
4	84	50.9	25	19 AAW71901	Dermatophagoides p
5	84	50.9	25	20 AAY50520	Dermatophagoides s
6	84	50.9	25	22 AAU19123	T-cell epitope con
7	84	50.9	26	14 AAR36424	DP11-20(1-26) a De
8	84	50.9	26	15 AAR51772	Der p II derived p
9	84	50.9	26	19 AAW72303	Dermatophagoides D
10	84	50.9	26	19 AAW71947	Dermatophagoides D

11	84	50.9	26	20 AAY50401	Dermatophagoides s
12	84	50.9	26	20 AAY50521	Dermatophagoides s
13	84	50.9	26	22 AAU19004	T-cell epitope con
14	84	50.9	26	22 AAU19124	T-cell epitope con
15	84	50.9	27	19 AAW72304	Dermatophagoides D
16	84	50.9	27	20 AAY50522	Dermatophagoides s
17	84	50.9	27	22 AAU19125	T-cell epitope con
18	83	50.3	26	14 AAR36426	DP11-20.2(1-26)S8,
19	83	50.3	26	15 AAR51774	Der p II derived p
20	83	50.3	26	19 AAW71949	Dermatophagoides D
21	83	50.3	26	20 AAY50403	Dermatophagoides s
22	83	50.3	26	22 AAU19006	T-cell epitope con
23	82	49.7	26	14 AAR36425	DP11-20.1(1-26)E8,
24	82	49.7	26	15 AAR51773	Der p II derived p
25	82	49.7	26	19 AAW71948	Dermatophagoides D
26	82	49.7	26	20 AAY50402	Dermatophagoides s
27	82	49.7	26	22 AAU19005	T-cell epitope con
28	81.5	49.4	28	23 AAU11422	Synthetic immunoge
29	79	47.9	14	21 AAY82633	Der pII B cell epi
30	79	47.9	16	14 AAR36433	DP11-2.1(11-26), a
31	79	47.9	16	15 AAR51781	Der p II derived p
32	79	47.9	16	19 AAW71956	Dermatophagoides D
33	79	47.9	16	20 AAY50410	Dermatophagoides s
34	79	47.9	16	22 AAU19013	T-cell epitope con
35	79	47.9	25	14 AAR36416	DP11-2(11-35) a De
36	79	47.9	25	16 AAR82550	Dermatophagoides p
37	79	47.9	25	19 AAW71939	Dermatophagoides D
38	79	47.9	25	20 AAY50393	Dermatophagoides s
39	79	47.9	25	22 AAU18996	T-cell epitope con
40	78	47.3	25	22 AAB49092	Amyloid beta/tetan
41	78	47.3	27	22 AAB49074	Amyloid beta/tetan
42	78	47.3	27	22 AAB49077	Amyloid beta/tetan
43	76	46.1	31	17 AAW06129	Anti-cholesteryl e
44	76	46.1	31	20 AAY02470	Fusion of a tetanu
45	76	46.1	31	23 ABB79179	Human cytomegalovi

#### ALIGNMENTS

```

RESULT 1
AAY82632
ID AAY82632 standard; peptide; 31 AA.
XX
AC AAY82632;
XX
DT 07-AUG-2000 (first entry)
XX
DE Tetanus toxoid T cell epitope and Der pII B cell epitope peptide.
XX
KW T cell epitope; B cell epitope; allergy; allergen; antigenic;
KW antiallergic; antiasthmatic; antiinflammatory; dermatological;
KW immunosuppressive; vaccine; rhinitis; sinusitis; bronchial asthma;
KW atopic dermatitis; acute urticaria; chronic urticaria;
KW gastro-intestinal syndrome; food allergen; oro-pharyngeal syndrome;
KW anaphylactic reaction; drug hypersensitivity; allergic reaction.
XX
OS Dermatophagoides pteronyssinus.
OS Clostridium tetani.
OS Synthetic.
XX
XX WO2000006694-A2.
XX
PD 10-FEB-2000.
XX
PF 20-JUL-1999; 99WO-BE00092.
XX
PR 30-JUL-1998; 98EP-0870167.
XX
PA (UNTO ) UCB SA.
XX
XX Saint-Remy J, Jacquemin M;
XX

```

DR WPI; 2000-422470/36.

XX New compound for prevention and treatment of allergies comprises at

PT least one allergen antigenic determinant recognized by a B cell and at

PT least one antigenic determinant which does not trigger T cell

PT activation -

XX

PS Claim 8; Page 35; 50pp; English.

XX

CC The present invention describes a compound (I) for the prevention and/or

CC treatment of allergy. The compound comprises at least one allergen

CC antigenic determinant (i) recognised by a B cell or an antibody secreted

CC by a B cell of a non-atopic individual and at least one antigenic

CC determinant (ii) different from the allergen that triggers T cell

CC activation. (i) has anti-allergic, antiasthmatic, antiinflammatory,

CC dermatological and immunosuppressive activities, and can be used in a

CC vaccine. (i) may be used in a pharmaceutical or cosmetic medicament to

CC treat and/or prevent allergies or a disease of allergic origin,

CC especially hypersensitivities. These include rhinitis, sinusitis,

CC bronchial asthma, atopic dermatitis, some forms of acute and chronic

CC urticaria, gastro-intestinal syndromes associated with the ingestion of

CC food allergens, oro-pharyngeal syndrome, anaphylactic reactions

CC associated with drug hypersensitivities and/or a mixture of these. The

CC use of (i) in the treatment of allergic conditions avoids the need for

CC drug treatment, which often causes undesirable side-effects. Also, prior

CC art drug therapies alleviate symptoms, but do not influence their

CC causes, however (i) actually combats the cause of an allergic reaction.

CC The present sequence represents a specifically claimed compound peptide

CC sequence from the present invention.

XX

SQ Sequence 31 AA;

Query Match 100.0%; Score 165; DB 21; Length 31;

Best Local Similarity 100.0%; Pred. No. 5.2e-19;

Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITELGGHEIKKVLVPGCHGS 31

DB 1 QYIKANSKFIGITELGGHEIKKVLVPGCHGS 31

RESULT 2

AAR83561

ID AAR83561 standard; peptide; 29 AA.

XX

AC AAR83561;

XX

DT 13-JUN-1996 (first entry)

XX

DE IgE CH4 region contg. peptide immunogen for treating allergies.

XX

KW IgE; CH4; immunoglobulin; epsilon; immunogen; helper T cell; epitope;

KW vaccine; allergy; antibody; constant heavy chain.

XX

OS Synthetic.

XX

PN W09526365-A1.

XX

PD 05-OCT-1995.

XX

PF 24-MAR-1995; 95WO-US03741.

XX

PR 25-OCT-1994; 94US-0328912.

PR

PR 28-MAR-1994; 94US-0218461.

XX

XX (UNBI-) UNITED BIOMEDICAL INC.

XX

PI Wang C;

XX

DR WPI; 1995-351297/45.

XX

PT Synthetic peptide-based immunogen contg. IgE CH4 peptide and helper

PT T cell epitope - useful for eliciting antibody prodn. for allergy

PT treatment

XX

PS Claim 5; Page 68-69; 87pp; English.

XX

CC AAR82592-R82600 and AAR83560-R83581 are peptide immunogens that are

CC useful in vaccines for treating allergic reactions. In the immunogens,

CC an IgE CH4 peptide is attached C-terminally to a series of amino acids

CC including a helper T cell epitope. The immunogen may also opt. contain

CC a fatty acid or fatty acid derivative, an invasin domain or alpha-NH2.

CC The immunogen produces high titres of antibodies to the effector site

CC in human IgE heavy chain (the CH4 domain peptide) which inhibit mast

CC cell activation and reduce allergen-induced IgE prodn. The immunogens

CC may be used in either a radially branching multimeric form or a

CC linearly arranged monomeric form.

XX

SQ Sequence 29 AA;

Query Match 52.1%; Score 86; DB 16; Length 29;

Best Local Similarity 100.0%; Pred. No. 2e-06;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITELGG 17

DB 3 QYIKANSKFIGITELGG 19

RESULT 3

AAR77122

ID AAR77122 standard; peptide; 25 AA.

XX

AC AAR77122;

XX

DT 30-MAY-1996 (first entry)

XX

DE Dermatophagoides pteronyssinus group II peptide DPII-20.9.

XX

KW House dust mite; DerpI; DerfI; DerpII; DerfII; allergen; allergy.

XX

OS Dermatophagoides pteronyssinus.

XX

PN W09528424-A1.

PN

PD 26-OCT-1995.

XX

PF 12-APR-1995; 95WO-US04481.

XX

PR 14-APR-1994; 94US-0227722.

XX

PA (IMMU-) IMMULOGIC PHARM CORP.

XX

PI Chen X, Evans S, Franzen HM, Kuo M, Shaked Z;

XX

DR WPI; 1995-373765/48.

XX

PT Compsns. contg. house mite allergen-derived peptide(s), some of

PT which are new - are used to treat allergy, and are stable, soluble

PT and able to induce T cell non-responsiveness

XX

PS Claim 1; Fig 1; 61pp; English.

XX

CC Claimed therapeutic compositions contain at least one of the peptides

CC DPI-21.2 and DPI-22.2 and also at least one of the new peptides

CC DPI-23.31, DPI-26.6, DPII-20.9, DPII-22.14 and DPII-25.15. The

CC compositions are useful for treating sensitivity to house dust mite

CC allergens. The peptides were identified by screening overlapping

CC peptides derived from D. pteronyssinus and D. farinae group I and II

CC allergens for T-cell reactivity in sensitised individuals.

CC The present sequence is that of novel peptide DPII-20.9.

XX

SQ Sequence 25 AA;

Query Match 50.9%; Score 84; DB 16; Length 25;

Best Local Similarity 63.6%; Pred. No. 3.5e-06;

Dermatophagoides sp major protein allergen D<sub>PTI</sub>-20.9.

Allergen; house dust mite; detection; sensitivity; T cell epitope; screening; allergic disorder; asthma; rhinitis; ectopic dermatitis; Der f I; Der p I; Der p II; Der f II.

Dermatophag  
US5968526-A.  
19-OCT-1999.

07-JUN-1995; 9505-04/85/Z.

19-MAY-1995; 9505-044530/.

I2-APK-1995;  
95WO-US04481.

Garman RD, Greenstein JL, Rogers BL, Franzen HM, Shaked Z, Chen X; Evans S, Kuo M:

Evans S, Kuo M;

WPI: 1999-590385/50.

Screening individuals for allergic reactions to T cell epitopes of major allergens from house dust mites -

individual is sensitive to Dermatophagoid

comprising combining a blood sample from

(DP II)) from Dermatophagoides (house dust)

may be used in (I). The sample and aller

polypeptides. The extent of binding is

individuals for sensitivity to Dermatoph

as asthma, rhinitis and eczopic dermatitis. AAY50360-Y50542 and AAY50546-Y50555 represent house dust mite allergen peptide fragments derived from Der p I, Der f II, Der f I and Der f II.

TGTTTGGHETKKVLTVPCHGS 31

3 VDKDCANHEIKKVLVPGCHGS 24

AAU19123 standard; Peptide; 25  
AAU19123:

04-DEC-2001 (first entry)  
T-cell epitope containing peptide DP11-20.9.

House dust mite; allergenic peptide; Der p 1

T-cell epitope.

*Dermatophagoides pteronyssinus*.



```

XX US6268491-B1.
PN
XX
XX
XX 31-JUL-2001.
XX
XX
XX 07-JUN-1995; 95US-0484296.
XX
XX 19-MAY-1995; 95US-0445307.
XX 16-OCT-1991; 91US-0777859.
XX 08-MAY-1992; 92US-0881396.
XX 14-APR-1993; 93WO-US03471.
XX 14-APR-1994; 94US-0227772.
XX
XX (IMMU-) IMMULOGIC PHARM CORP.
XX
XX Garman RD, Greenstein JL, Kuo M, Rogers BL, Franzen HM, Chen X;
PI Evans S, Shaked Z;
XX
XX WPI; 2001-549074/61.
XX
XX Peptides comprising T cell groups of the major allergens from
PT Dermatophagoides (house dust mites), useful for treating house dust
PT mite allergy in humans, and for diagnosing sensitivity to house dust
PT mite protein allergens -
XX
XX Claim 2; Figure 30; 158pp; English.
XX
XX The invention relates to an isolated peptide of the major protein
CC allergens of the genus Dermatophagoides, which comprises at least one T
CC cell group of a protein allergen from Der p (DP) I, DP II, Der f (DF) I
CC or DF II. The isolated peptide comprises at least two regions,
CC each region comprising at least one T cell group of a protein allergen
CC of the genus Dermatophagoides. The regions are derived from the same or
CC different protein allergens of the genus Dermatophagoides. The peptides
CC are useful for treating house dust mite allergy in humans. The peptides
CC are also useful for detecting or diagnosing sensitivity to house dust
CC mite protein allergens. The present peptides have similar or enhanced
CC therapeutic properties as the naturally-occurring allergen, but have
CC reduced side effects, and increased solubility and stability. The
CC present sequence represents an allergenic T-cell epitope containing
CC peptide derived from the Dermatophagoides allergenic proteins.
XX
XX Sequence 25 AA;
SQ
Query Match 50.9%; Score 84; DB 22; Length 25;
Best Local Similarity 63.6%; Pred. No. 3.5e-06;
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 IGITELGCGHEIKKVLVPGCHGS 31
: : : |||||
Db 3 VDVKDCANHEIKKVLVPGCHGS 24

RESULT 7
AAR36424
ID AAR36424 standard; peptide; 26 AA.
XX
XX AAR36424;
AC
XX
XX 12-AUG-1993 (first entry)
XX
XX DPTI-20(1-26) a Dermatophagoides protein allergen.
XX
XX T cell epitope; house dust mite; allergy; soluble; Der pII.
XX
XX Synthetic.
XX
XX WO9308279-A.
XX
XX 25-APR-1993.
XX
XX 15-OCT-1992; 92WO-US08637.
XX

```

```

PR 16-OCT-1991; 91US-0777859.
PR 08-MAY-1992; 92US-0881396.
XX
XX (IMMU-) IMMULOGIC PHARM CORP.
XX
XX Garman RD, Greenstein JL, Kuo MC, Rogers BL;
PI
XX WPI; 1993-152472/18.
XX
XX Isolated peptide(s) of Dermatophagoides protein allergens - for
PT diagnosis and treatment of sensitivity to house dust mite
PT
XX
XX Claim 44; Fig 3; 176pp; English.
XX
XX The peptide is one of a series of overlapping peptides synthesised by
CC standard techniques to cover the whole Dermatophagoides
CC pteronyssinus Der pII sequence. The T cell epitopes of the protein
CC were mapped by detection of the peptide's ability to stimulate T
CC cell activity. The peptides may be used for diagnosis and treatment
CC of sensitivity to house dust mite allergens. When administered to
CC house dust mite sensitive individuals, the peptides are capable of
CC modifying the allergic response to the allergens. The peptides may
CC be modified for e.g. increasing solubility, enhancing therapeutic or
CC preventive efficacy or stability.
CC See also AAR34686-700 and AAR36398-490.
XX
XX Sequence 26 AA;
SQ
Query Match 50.9%; Score 84; DB 14; Length 26;
Best Local Similarity 63.6%; Pred. No. 3.7e-06;
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 IGITELGCGHEIKKVLVPGCHGS 31
: : : |||||
Db 3 VDVKDCANHEIKKVLVPGCHGS 24

RESULT 8
AAR51772
ID AAR51772 standard; Protein; 26 AA.
XX
XX AAR51772;
AC
XX
XX 01-FEB-1995 (first entry)
XX
XX Der p II derived peptide, DP II-20(1-26).
XX
XX Group I; protein allergen; house dust mite; D. pteronyssinus; Der p I;
KW homology; D. farinae; Der f I; group II; Der p II; Der f II; T-cell;
KW epitopes; fusion peptides; antigenic fragments; substitution; deletion;
KW addition; chemical synthesis; chemical cleavage; recombinant techniques;
KW allergic response; immunoglobulin E; IgE; immunotherapy; anaphylaxis;
KW IgE-mediated responses; anergise; lymphokine secretion profile; modify;
KW T cell subpopulations; unresponsive; immune response; tolerance.
XX
XX Dermatophagoides pteronyssinus.
OS
XX
XX 2A9302677-A.
XX
XX 26-JAN-1994.
XX
XX 16-APR-1993; 93ZA-0002677.
XX
XX 16-APR-1993; 93ZA-0002677.
XX
XX (IMMU-) IMMULOGIC PHARM CORP.
XX
XX Garman RD, Greenstein JL, Kuo M, Rogers BL;
PI
XX WPI; 1994-126807/15.
XX
XX Isolated and/or modified peptides comprising T-cell epitopes - of
PT major protein allergens of genus Dermatophagoides, used to treat

```

PT or diagnose sensitivity to house dust mites  
 XX Claim 28; Page 70; 154pp; English.  
 PS  
 XX The sequences given in AAR51731-841 represent T-cell epitopes derived from the group I and II protein allergens from the house dust mite D. farinae and D. pteronyssinus, Der f I, Der f II, Der p I and Der p II respectively. The Der f II proteinsCC shows high homology having an identity of 88%, with an identity of 81% between the two group I proteins (see also AAR51721-30). Fusion peptides may be produced which comprise at least two or these antigenic fragments. Each region of these fusion peptides may be derived from the same, or different, mite allergens. The antigenic fragments may be altered by substitution, deletion or addition to enhance their antigenicity. These peptides may be produced by chemical synthesis, chemical cleavage of the protein allergen or by recombinant techniques. These peptides, or the fusion peptides, when administered to a house dust mite sensitive individual, are capable of modifying the allergic response of the individual to the allergen. The peptides do not bind to immunoglobulin E (IgE), or bind IgE to a lesser extent than the full length protein allergen. This reduces the major complications of standard immunotherapy, which are IgE-mediated responses such as anaphylaxis. Exposure of mite allergic patients to these peptides may tolerate or anergise appropriate T cell subpopulations such that they become unresponsive to mite allergens and do not participate in mounting an immune response upon exposure. Administration of the peptides may also modify the lymphokine secretion profile as compared with exposure to the naturally occurring mite protein allergen.

XX  
 SQ Sequence 26 AA;  
 Query Match 50.9%; Score 84; DB 15; Length 26;  
 Best Local Similarity 63.6%; Pred. No. 3.7e-06;  
 Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31  
 : : |||||  
 Db 3 VDVKDCANHEIKKVLVPGCHGS 24

RESULT 9  
 AAW72303  
 ID AAW72303 standard; peptide; 26 AA.  
 XX  
 AC AAW72303;  
 XX  
 DT 16-DEC-1998 (first entry)  
 XX  
 DE Dermatophagoides Der p II protein peptide DP11-20.10.  
 XX  
 KW genus Dermatophagoides; major protein allergen; T cell epitope;  
 KW Der p I; Der p II; Der f I; Der f II; house dust mite allergy.  
 XX  
 OS Dermatophagoides sp.  
 XX  
 PN US5820862-A.  
 XX  
 PD 13-OCT-1998.  
 XX  
 PF 07-JUN-1995; 95US-0482142.  
 XX  
 PR 19-MAY-1995; 95US-0445307.  
 PR 14-APR-1994; 94US-0227772.  
 PR 07-JUN-1995; 95US-0482142.  
 XX  
 PA (IMMU-) IMMULOGIC PHARM CORP.  
 XX  
 PI Chen X, Evans S, Franzen HM, Garman RD, Greenstein JL;  
 PI Kuo M, Rogers BL, Shaked Z;  
 XX  
 DR WPI; 1998-567590/48.  
 XX  
 PT Dermatophagoides allergen peptides - useful for treating house dust

PT mite allergy  
 XX  
 PS Disclosure; Column 165-166; 155pp; English.  
 XX The present invention describes peptides for treating sensitivity to house dust mite allergens from the genus Dermatophagoides. Peptides within the scope of the invention comprise at least one T cell epitope, or preferably at least two T cell epitopes of a protein allergen selected from the allergens Der p I, Der p II, Der f I, or Der f II. The invention also describes modified peptides having similar or enhanced therapeutic properties as the corresponding, naturally occurring allergen, but having reduced side effects. AAW71912 to CC AAW72000, and AAW72257 to AAW72330 represent peptides from the present CC invention.

XX  
 SQ Sequence 26 AA;  
 Query Match 50.9%; Score 84; DB 19; Length 26;  
 Best Local Similarity 63.6%; Pred. No. 3.7e-06;  
 Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31  
 : : |||||  
 Db 3 VDVKDCANHEIKKVLVPGCHGS 24

RESULT 10  
 AAW71947  
 ID AAW71947 standard; peptide; 26 AA.  
 XX  
 AC AAW71947;  
 XX  
 DT 16-DEC-1998 (first entry)  
 XX  
 DE Dermatophagoides Der p II protein peptide DP11-20.  
 XX  
 KW genus Dermatophagoides; major protein allergen; T cell epitope;  
 KW Der p I; Der p II; Der f I; Der f II; house dust mite allergy.  
 XX  
 OS Dermatophagoides sp.  
 XX  
 PN US5820862-A.  
 XX  
 PD 13-OCT-1998.  
 XX  
 PF 07-JUN-1995; 95US-0482142.  
 XX  
 PR 19-MAY-1995; 95US-0445307.  
 PR 14-APR-1994; 94US-0227772.  
 PR 07-JUN-1995; 95US-0482142.  
 XX  
 PA (IMMU-) IMMULOGIC PHARM CORP.  
 XX  
 PI Chen X, Evans S, Franzen HM, Garman RD, Greenstein JL;  
 PI Kuo M, Rogers BL, Shaked Z;  
 XX  
 DR WPI; 1998-567590/48.  
 XX  
 PT Dermatophagoides allergen peptides - useful for treating house dust

PT mite allergy  
 XX  
 PS Disclosure; Column 165-166; 155pp; English.  
 XX The present invention describes peptides for treating sensitivity to house dust mite allergens from the genus Dermatophagoides. Peptides within the scope of the invention comprise at least one T cell epitope, or preferably at least two T cell epitopes of a protein allergen selected from the allergens Der p I, Der p II, Der f I, or Der f II. The invention also describes modified peptides having similar or enhanced therapeutic properties as the corresponding, naturally occurring allergen, but having reduced side effects. AAW71912 to CC AAW72000, and AAW72257 to AAW72330 represent peptides from the present CC invention.

XX SQ Sequence 26 AA; Query Match 50.9%; Score 84; DB 19; Length 26;  
Best Local Similarity 63.6%; Pred. No. 3.7e-06;  
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 10 IGITELGGHEIKKVLPGCHGS 31  
: : : |||||  
Db 3 VDVKDCANHEIKKVLPGCHGS 24

RESULT 11  
AAV50401 ID AAY50401 standard; Peptide; 26 AA.  
XX AC AAY50401;  
XX XX  
DT 25-JAN-2000 (first entry)  
XX Dermatophagoides sp major protein allergen DP II-20.  
DE XX  
KW Allergen; house dust mite; detection; sensitivity: T cell epitope;  
KW screening; allergic disorder; asthma; rhinitis; ectopic dermatitis;  
KW Der f I; Der p I; Der p II; Der f II.  
XX OS Dermatophagoides sp.  
XX US5968526-A.  
PN XX  
PD 19-OCT-1999.  
XX XX  
PF 07-JUN-1995; 95US-0478572.  
XX PR 19-MAY-1995; 95US-0445307.  
PR 14-APR-1994; 94US-0227772.  
PR 12-APR-1995; 95WO-US04481.  
XX XX  
PA (IMMU-) IMMULOGIC PHARM CORP.  
XX XX  
PI Garman RD, Greenstein JL, Rogers BL, Franzen HM, Shaked Z, Chen X;  
PI Evans S, Kuo M;  
DR WPI; 1999-590385/50.  
XX Screening individuals for allergic reactions to T cell epitopes of  
PT major allergens from house dust mites -  
XX Claim 3o; Column 105-106; 158pp; English.

This invention describes a novel method (I) for detecting whether an individual is sensitive to Dermatophagoides (house dust mites). The method involves detecting sensitivity to house dust mites in patients, comprising combining a blood sample from the individual with 1 or more isolated T cell epitopes of the protein allergens I and II ((DP I) and (DP II)) from Dermatophagoides (house dust mites). 32 T cell epitopes with varying, defined amino acids sequences (given in the specification) may be used in (I). The sample and allergens are combined under conditions appropriate for the binding of blood components with the polypeptides. The extent of binding is then indicative of the sensitivity of the patient to house dust mites. (I) may be used to screen individuals for sensitivity to Dermatophagoides (house dust mites). The house dust mite is a major cause of a variety of allergic disorders such as asthma, rhinitis and ectopic dermatitis. AAY50360-Y50542 and CC AAY5046-fY50555 represent house dust mite allergen peptide fragments derived from Der p I, Der f II, Der f I and Der f II.  
XX XX

SQ Sequence 26 AA; Query Match 50.9%; Score 84; DB 20; Length 26;  
Best Local Similarity 63.6%; Pred. No. 3.7e-06;  
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

XX	T-cell epitope containing peptide DPPII-20.10.
DE	
XX	House dust mite; allergenic peptide; Der p I; Der p II; Der f I;
KW	Der f II; antiallergenic; immunostimulant; house dust mite allergy;
KW	T-cell epitope.
XX	
OS	Dermatophagoides pteronyssinus.
XX	
PN	US6268491-B1.
XX	
PD	31-JUL-2001.
XX	
PF	07-JUN-1995; 95US-0484296.
XX	
PR	19-MAY-1995; 95US-0445307.
PR	16-OCT-1991; 91US-0777859.
PR	08-MAY-1992; 92US-0881396.
PR	14-APR-1993; 93WO-US03471.
PR	14-APR-1994; 94US-0227772.
XX	
PA	(IMMU-) IMMULOGIC PHARM CORP.
XX	
PI	Garman RD, Greenstein JL, Kuo M, Rogers BL, Franzen HM, Chen X;
PI	Evans S, Shaked Z;
XX	
DR	WPI; 2001-549074/61.
XX	
PT	Peptides comprising T cell groups of the major allergens from
PT	Dermatophagoides (house dust mites), useful for treating house dust
PT	mite allergy in humans, and for diagnosing sensitivity to house dust
PT	mite protein allergens -
PS	
PS	Claim 2; Figure 30; 158pp; English.
XX	
CC	The invention relates to an isolated peptide of the major protein
CC	allergens of the genus Dermatophagoides, which comprises at least one T
CC	cell group of a protein allergen from Der p (DP) I, DP II, Der f (DF) I
CC	or DF II. The isolated peptide comprises at least two regions,
CC	each region comprising at least one T cell group of a protein allergen
CC	of the genus Dermatophagoides. The regions are derived from the same or
CC	different protein allergens of the genus Dermatophagoides. The peptides
CC	are useful for treating house dust mite allergy in humans. The peptides
CC	are also useful for detecting or diagnosing sensitivity to house dust
CC	mite protein allergens. The present peptides have similar or enhanced
CC	therapeutic properties as the naturally-occurring allergen, but have
CC	reduced side effects, and increased solubility and stability. The
CC	present sequence represents an allergenic T-cell epitope containing
CC	peptide derived from the Dermatophagoides allergenic proteins.
XX	
SQ	Sequence 26 AA:
	Query Match 50.9%; Score 84; DB 22; Length 26;
	Best Local Similarity 63.6%; Pred. No. 3.7e-06;
	Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
Oy	10 IGITELGGHEIKKVLPGCHGS 31 : : :
Db	3 VDVKCANHEIKKVLPGCHGS 24 : : :
RESULT 15	
AAW72304	
ID	AAW72304 standard; peptide; 27 AA.
XX	
AC	AAW72304;
XX	
DT	16-DEC-1998 (first entry)
XX	
DE	Dermatophagoides Der p II protein peptide DPPII-20.8.
XX	
KW	genus Dermatophagoides; major protein allergen; T cell epitope;
KW	Der p I; Der p II; Der f I; Der f II; house dust mite allergy.

```

XX OS Dermatophagoides sp.
XX PN US5820862-A.
XX PD 13-OCT-1998.
XX PF 07-JUN-1995; 95US-0482142.
XX PR 19-MAY-1995; 95US-0445307.
XX PR 14-APR-1994; 94US-0227772.
XX PR 07-JUN-1995; 95US-0482142.
XX PA (IMMU-) IMMULOGIC PHARM CORP.
XX PI Chen X, Evans S, Franzen HM, Garman RD, Greenstein JL;
XX PI Kuo M, Rogers BL, Shaked Z;
XX DR WPI; 1998-567590/48.
XX PT Dermatophagoides allergen peptides - useful for treating house dust
XX PT mite allergy
XX PS Disclosure; Column 165-166; 155pp; English.
XX CC The present invention describes peptides for treating sensitivity to
XX CC house dust mite allergens from the genus Dermatophagoides. Peptides
XX CC within the scope of the invention comprise at least one T cell epitope,
XX CC or preferably at least two T cell epitopes of a protein allergen
XX CC selected from the allergens Der p I, Der p II, Der f I, or Der f II.
XX CC The invention also describes modified peptides having similar or
XX CC enhanced therapeutic properties as the corresponding, naturally
XX CC occurring allergen, but having reduced side effects. AAW71912 to
XX CC AAW72000, and AAW72257 to AAW72330 represent peptides from the present
XX CC invention.
XX SQ Sequence 27 AA;
    Query Match 50.9%; Score 84; DB 19; Length 27;
    Best Local Similarity 63.6%; Pred. No. 3.8e-06;
    Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
    QY 10 IGITELGGHEIKKVLVPGCHGS 31
    Db 3 VDVKDCANHEIKKVLVPGCHGS 24
    Search completed: December 4, 2002, 13:06:09
    Job time : 34 secs

```

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 13:07:06 ; Search time 10 Seconds  
(without alignments)  
50.351 Million cell updates/sec

Title: US-09-362-731a-1  
Perfect score: 165  
Sequence: 1 QYKANSKFGITELGGHEIKVLVPCCHGS 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 40221

Minimum DB seq length: 0  
Maximum DB seq length: 31

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81.5	49.4	28	10	US-09-848-834A-11
2	78.5	47.6	29	10	US-09-732-754-1
3	76	46.1	31	10	US-09-943-548-2
4	76	46.1	31	10	US-09-983-019-5
5	74	44.8	15	10	US-09-862-849-2
6	74	44.8	16	10	US-09-848-834A-2
7	74	44.8	31	10	US-09-848-834A-15
8	70	42.4	29	10	US-09-983-019-8
9	70	42.4	29	10	US-09-983-019-9
10	70	42.4	31	10	US-09-983-019-3
11	70	42.4	31	10	US-09-983-019-6
12	63	38.2	22	10	US-09-860-793-7
13	58.5	35.5	14	9	US-10-044-034-21
14	54	32.7	28	10	US-09-864-761-47156
15	36	21.8	12	10	US-09-784-887B-6
16	33	20.0	23	10	US-09-030-619-231
17	33	20.0	28	9	US-09-749-637A-310
18	33	20.0	30	9	US-09-749-637A-92
19	32.5	19.7	30	10	US-09-864-761-36656

20	32	19.4	25	10	US-09-864-761-42617
21	32	19.4	30	9	US-09-747-419-34
22	31	18.8	15	10	US-09-826-752-20
23	31	18.8	25	10	US-09-821-984-26
24	31	18.8	31	9	US-09-956-206A-32
25	31	18.8	31	10	US-09-864-761-37570
26	30.5	18.5	26	10	US-09-860-793-6
27	30	18.2	26	10	US-09-864-761-35140
28	30	18.2	26	10	US-09-867-852-151
29	30	18.2	27	10	US-09-864-761-48517
30	30	18.2	28	10	US-09-728-721-18
31	30	18.2	28	12	US-10-105-931-18
32	29	17.6	9	10	US-09-765-086-193
33	29	17.6	15	10	US-09-748-739A-9
34	29	17.6	20	10	US-09-726-643-83
35	29	17.6	21	9	US-09-349-755-13
36	29	17.6	21	9	US-09-349-755-20
37	29	17.6	21	9	US-09-349-755-39
38	29	17.6	21	9	US-09-166-334-13
39	29	17.6	21	9	US-09-166-334-20
40	29	17.6	21	9	US-09-166-334-39
41	29	17.6	21	10	US-09-350-206-13
42	29	17.6	21	10	US-09-350-206-20
43	29	17.6	21	10	US-09-350-206-39
44	29	17.6	22	10	US-09-864-761-42066
45	29	17.6	22	10	US-09-953-510-34

## ALIGNMENTS

## RESULT 1

US-09-848-834A-11  
; Sequence 11, Application US/09848834A  
; Patent No. US20020076416A1  
; GENERAL INFORMATION:  
; APPLICANT: Aphton Corporation  
; TITLE OF INVENTION: Chimeric Peptide Immunogens  
; FILE REFERENCE: 1102865-0047  
; CURRENT APPLICATION NUMBER: US/09/848,834A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/202,328  
; PRIOR FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 11  
; LENGTH: 28  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 830-844 of  
; OTHER INFORMATION: Tetanus toxoid precursor (Tentoxylisin) linked by a spacer to  
; OTHER INFORMATION: ino acid sequence 2-10 of the GnRH hormone  
; NAME/KEY: MOD.RES  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Amidated-glutamine  
; NAME/KEY: MOD.RES  
; LOCATION: (28)..(28)  
; OTHER INFORMATION: Amidated-glycine or glycineamide  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(15)  
; OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus Toxoid Precursor  
; OTHER INFORMATION: (tentoxylisin)  
; NAME/KEY: PEPTIDE  
; LOCATION: (16)..(19)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (20)..(28)  
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone  
US-09-848-834A-11

Query Match 49.4%; Score 81.5; DB 10; Length 28;  
Best Local Similarity 61.3%; Pred. No. 3e-06;

```
Matches 19; Conservative 0; Mismatches 3; Indels 9; Gaps 1;

QY 1 QYIKANSKFIGITELGGHEIKKVLVPGCHGS 31
| | | | | | | | | | | | | | | | | | | | |
Db 1 QYIKANSKFIGITELG-----PSLHWS 22

RESULT 2
US-09-732-754-1
; Sequence 1, Application US/09732754
; Patent No. US20020031523A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; APPLICANT: GRAS-MASSE, HELENE
; APPLICANT: BENMOMAMED, LLACHIR
; TITLE OF INVENTION: SYSTEMIC IMMUNE RESPONSE INDUCED BY MUCOSAL ADMINISTRATION OF LIP
; TITLE OF INVENTION: TAILED POLYPEPTIDES WITHOUT ADJUVANT
; FILE REFERENCE: 2008050855
; CURRENT APPLICATION NUMBER: US/09/732,754
; CURRENT FILING DATE: 2000-12-11
; PRIOR APPLICATION NUMBER: US 60/169,952
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Polypeptide
; NAME/KEY: LIPID
; LOCATION: (1)..(1)
; OTHER INFORMATION: N-epsilon Pam
US-09-732-754-1

Query Match 47.6%; Score 78.5; DB 10; Length 29;
Best Local Similarity 60.0%; Pred. No. 8.9e-06;
Matches 18; Conservative 2; Mismatches 5; Indels 5; Gaps 1;

QY 1 QYIKANSKFIGITELGGHEIKKVLVPGCHG 30
| | | | | | | | | | | | | | | | | | | | |
Db 4 QYIKANSKFIGITERG-----RILKEPVHG 28

RESULT 3
US-09-943-548-2
; Sequence 2, Application US/09943548
; Patent No. US20020042364A1
; GENERAL INFORMATION:
; APPLICANT: Rittershaus, Charles W.
; APPLICANT: Thomas, Lawrence J.
; TITLE OF INVENTION: MODULATION OF CHOLESTERYL ESTER TRANSFER PROTEIN (CETP) ACTIVITY
; FILE REFERENCE: TCS-411.1P US-1; TCS-411.1P US-2
; CURRENT APPLICATION NUMBER: US/09/943,548
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 08/432,483
; PRIOR FILING DATE: 1995-05-01
; PRIOR APPLICATION NUMBER: PCT/US96/06147
; PRIOR FILING DATE: 1996-05-01
; PRIOR APPLICATION NUMBER: 08/945,289
; PRIOR FILING DATE: 1997-10-17
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: vaccine peptide of the invention
US-09-943-548-2

Query Match 46.1%; Score 76; DB 10; Length 31;
Best Local Similarity 93.8%; Pred. No. 2.3e-05;
```

```
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITELG 16
| | | | | | | | | | | | | | | | |
Db 2 QYIKANSKFIGITERG 17

RESULT 4
US-09-983-019-5
; Sequence 5, Application US/09983019
; Patent No. US20020146820A1
; GENERAL INFORMATION:
; APPLICANT: Diamond, Don J.
; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS P
; FILE REFERENCE: 1954-347US
; CURRENT APPLICATION NUMBER: US/09/983,019
; CURRENT FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/241,944
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: ()..( )
; OTHER INFORMATION: Human cytomegalovirus vaccine peptide
; NAME/KEY: LIPID
; LOCATION: (1)..(1)
; OTHER INFORMATION: palmitic acid
US-09-983-019-5

Query Match 46.1%; Score 76; DB 10; Length 31;
Best Local Similarity 59.3%; Pred. No. 2.3e-05;
Matches 16; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITELGGHEIKKVLVPG 27
| | | | | | | | | | | | | | | | | | |
Db 4 QYIKANSKFIGITEAAQAQTVTSTPVG 30

RESULT 5
US-09-862-849-2
; Sequence 2, Application US/09862849
; Patent No. US20020013274A1
; GENERAL INFORMATION:
; APPLICANT: Sudhir Paul
; APPLICANT: Larry J. Smith
; APPLICANT: Gennady Gololobov
; TITLE OF INVENTION: Methods for Identifying Inducers and Inhibitors of Proteolytic
; FILE REFERENCE: UNMC 63123 DIV
; CURRENT APPLICATION NUMBER: US/09/862,849
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: US 09/046,373
; PRIOR FILING DATE: 1998-03-23
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-09-862-849-2

Query Match 44.8%; Score 74; DB 10; Length 15;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15
| | | | | | | | | | | | |
Db 1 QYIKANSKFIGITEL 15
```

Query Match	44.8%;	Score 74;	DB 10;	Length 31;
Best Local Similarity	100.0%;	Pred. No. 4.7e-05;		
Matches	15;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
Qy	1 QYIKANSKFIGITEL 15			
Db	17 QYIKANSKFIGITEL 31			
RESULT 8				
US-09-983-019-8				
; Sequence 8, Application US/09983019				
; Patent No. US20020146820A1				
; GENERAL INFORMATION:				
; APPLICANT: Diamond, Don J.				
; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS P				
; FILE REFERENCE: 1954-347US				
; CURRENT APPLICATION NUMBER: US/09/983,019				
; CURRENT FILING DATE: 2001-10-22				
; PRIOR APPLICATION NUMBER: US 60/241,944				
; PRIOR FILING DATE: 2000-10-20				
; NUMBER OF SEQ ID NOS: 14				
; SOFTWARE: PatentIn version 3.0				
; SEQ ID NO 8				
; LENGTH: 29				
; TYPE: PRT				
; ORGANISM: Artificial Sequence				
; FEATURE:				
; NAME/KEY: misc_feature				
; LOCATION: ()..()				
; OTHER INFORMATION: Human cytomegalovirus vaccine peptide				
US-09-983-019-8				
Query Match	42.4%;	Score 70;	DB 10;	Length 29;
Best Local Similarity	100.0%;	Pred. No. 0.00018;		
Matches	14;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
Qy	1 QYIKANSKFIGITE 14			
Db	16 QYIKANSKFIGITE 29			
RESULT 9				
US-09-983-019-9				
; Sequence 9, Application US/09983019				
; Patent No. US20020146820A1				
; GENERAL INFORMATION:				
; APPLICANT: Diamond, Don J.				
; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS P				
; FILE REFERENCE: 1954-347US				
; CURRENT APPLICATION NUMBER: US/09/983,019				
; CURRENT FILING DATE: 2001-10-22				
; PRIOR APPLICATION NUMBER: US 60/241,944				
; PRIOR FILING DATE: 2000-10-20				
; NUMBER OF SEQ ID NOS: 14				
; SOFTWARE: PatentIn version 3.0				
; SEQ ID NO 9				
; LENGTH: 29				
; TYPE: PRT				
; ORGANISM: Artificial Sequence				
; FEATURE:				
; NAME/KEY: misc_feature				
; LOCATION: ()..()				
; OTHER INFORMATION: Human cytomegalovirus vaccine peptide				
US-09-983-019-9				
Query Match	42.4%;	Score 70;	DB 10;	Length 29;
Best Local Similarity	100.0%;	Pred. No. 0.00018;		
Matches	14;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
Qy	1 QYIKANSKFIGITE 14			
Db	1 QYIKANSKFIGITE 14			



```
RESULT 10
US-09-983-019-3
; Sequence 3, Application US/09983019
; Patent No. US20020146820A1
; GENERAL INFORMATION:
; APPLICANT: Diamond, Don J.
; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS PP15
; FILE REFERENCE: 1954-347US
; CURRENT APPLICATION NUMBER: US/09/983,019
; CURRENT FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/241,944
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Human cytomegalovirus vaccine peptide
; NAME/KEY: LIPID
; LOCATION: (1)..(1)
; OTHER INFORMATION: di-palmitic acid
US-09-983-019-3

Query Match 42.4%; Score 70; DB 10; Length 31;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITE 14
Db 4 QYIKANSKFIGITE 17

RESULT 11
US-09-983-019-6
; Sequence 6, Application US/09983019
; Patent No. US20020146820A1
; GENERAL INFORMATION:
; APPLICANT: Diamond, Don J.
; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS PP15
; FILE REFERENCE: 1954-347US
; CURRENT APPLICATION NUMBER: US/09/983,019
; CURRENT FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/241,944
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Human cytomegalovirus vaccine peptide
; NAME/KEY: LIPID
; LOCATION: (1)..(1)
; OTHER INFORMATION: palmitic acid
US-09-983-019-6

Query Match 42.4%; Score 70; DB 10; Length 31;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITE 14
Db 18 QYIKANSKFIGITE 31
```

```
RESULT 12
US-09-860-793-7
; Sequence 7, Application US/09860793
; Patent No. US20020136734A1
; GENERAL INFORMATION:
; APPLICANT: Pruett, John H
; APPLICANT: Temeyer, Kevin B
; APPLICANT: Kunz, Sidney E
; APPLICANT: Fisher, William F
; TITLE OF INVENTION: Vaccines for the Protection of Cattle from Psoroptic
; FILE REFERENCE: Docket 0047.96 - John H. Pruett et al.
; CURRENT APPLICATION NUMBER: US/09/860,793
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/366,603
; PRIOR FILING DATE: 1999-08-03
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-09-860-793-7

Query Match 38.2%; Score 63; DB 10; Length 22;
Best Local Similarity 47.6%; Pred. No. 0.0015;
Matches 10; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHG 30
Db 2 VDVKDCANNEIKKVMVDGCHG 22

RESULT 13
US-10-044-034-21
; Sequence 21, Application US/10044034
; Patent No. US20020169264A1
; GENERAL INFORMATION:
; APPLICANT: JACKSON, DAVID C.
; APPLICANT: O'BRIEN-SIMPSON, NEIL M.
; APPLICANT: BROWN, LORENA E.
; APPLICANT: EDE, NICHOLAS J.
; APPLICANT: BRANDT, EVELYN R.
; APPLICANT: GOOD, MICHAEL F.
; TITLE OF INVENTION: POLYMERS INCORPORATING PEPTIDES
; FILE REFERENCE: FBRC:006
; CURRENT APPLICATION NUMBER: US/10/044,034
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: P05071
; PRIOR FILING DATE: 1997-02-11
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-044-034-21

Query Match 35.5%; Score 58.5; DB 9; Length 14;
Best Local Similarity 93.3%; Pred. No. 0.0044;
Matches 14; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 QYIKANSKFIGITEL 15
Db 1 QYIKANS-FIGITEL 14

RESULT 14
US-09-864-761-47156
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 13:05:35 ; Search time 15 Seconds  
(without alignments)  
60.807 Million cell updates/sec

Title: US-09-362-731A-1

Perfect score: 165

Sequence: 1 QYIKANSKFIGITELGGHEIKKVLVPCGCHS 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 159827

Minimum DB seq length: 0  
Maximum DB seq length: 31

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. Is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84	50.9	25	2	US-08-482-142-169
2	84	50.9	25	2	US-08-478-572-169
3	84	50.9	25	4	US-08-484-296-169
4	84	50.9	25	5	PCT-US95-04481-31
5	84	50.9	26	2	US-08-482-142-50
6	84	50.9	26	2	US-08-482-142-170
7	84	50.9	26	2	US-08-478-572-50
8	84	50.9	26	2	US-08-478-572-170
9	84	50.9	26	4	US-08-484-296-50
10	84	50.9	26	4	US-08-484-296-170
11	84	50.9	27	2	US-08-482-142-171
12	84	50.9	27	2	US-08-478-572-171
13	84	50.9	27	4	US-08-484-296-171
14	83	50.3	26	2	US-08-482-142-52
15	83	50.3	26	2	US-08-478-572-52
16	83	50.3	26	2	US-08-484-296-52
17	82	49.7	26	2	US-08-482-142-51
18	82	49.7	26	2	US-08-478-572-51
19	82	49.7	26	4	US-08-484-296-51
20	79	47.9	16	2	US-08-482-142-59
21	79	47.9	16	2	US-08-478-572-59
22	79	47.9	16	4	US-08-484-296-59
23	79	47.9	25	2	US-08-482-142-42
24	79	47.9	25	2	US-08-478-572-42
25	79	47.9	25	4	US-08-484-296-42
26	79	47.9	25	5	PCT-US95-04481-19
27	76	46.1	31	4	US-08-432-483A-2

28 75 45.5 31 5 PCT-US93-11703-63  
29 74 44.8 15 2 US-08-319-704-10  
30 74 44.8 15 2 US-08-661-052-6  
31 74 44.8 15 2 US-08-460-502-7  
32 74 44.8 15 4 US-09-046-373-2  
33 74 44.8 15 4 US-09-188-082-6  
34 74 44.8 15 4 US-09-364-088-6  
35 74 44.8 15 4 US-09-102-716-6  
36 74 44.8 15 4 US-09-148-711A-7  
37 74 44.8 15 4 US-08-716-249-4  
38 74 44.8 15 5 PCT-US93-11703-69  
39 74 44.8 16 4 US-09-248-588-55  
40 74 44.8 17 1 US-08-446-692-4  
41 74 44.8 17 2 US-08-488-351A-4  
42 74 44.8 17 3 US-09-100-409A-40  
43 74 44.8 17 5 PCT-US95-08596-23  
44 74 44.8 17 5 PCT-US95-13841-7  
45 74 44.8 26 2 US-08-482-142-54

#### ALIGNMENTS

RESULT 1  
US-08-482-142-169  
; Sequence 169, Application US/08482142  
; Patent No. 5820862  
; GENERAL INFORMATION:  
; APPLICANT: Garman, Richard  
; APPLICANT: Greenstein, Julia  
; APPLICANT: Kuo, Mei-chang  
; APPLICANT: Rogers, Bruce  
; APPLICANT: Franzen, Henry  
; APPLICANT: Chen, Xian  
; APPLICANT: Evans, Sean  
; APPLICANT: Shaked, Ze'ev  
; TITLE OF INVENTION: T CELL EPTOPES OF THE MAJOR ALLERGENS  
; NUMBER OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
; NUMBER OF SEQUENCES: 207  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
; STREET: 610 LINCOLN STREET  
; CITY: WALTHAM  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/482,142  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION NUMBER: US/08/445,307  
; FILING DATE: 07 June 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CRAIG, ANNE I.  
; REGISTRATION NUMBER: 32,976  
; REFERENCE/DOCKET NUMBER: 017.60S  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 466-6000  
; TELEFAX: (617) 466-6040  
; INFORMATION FOR SEQ ID NO: 169:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 25 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: internal  
US-08-482-142-169

Sequence 63, Appl  
Sequence 10, Appl  
Sequence 6, Appl  
Sequence 7, Appl  
Sequence 2, Appl  
Sequence 6, Appl  
Sequence 6, Appl  
Sequence 6, Appl  
Sequence 7, Appl  
Sequence 4, Appl  
Sequence 55, Appl  
Sequence 4, Appl  
Sequence 4, Appl  
Sequence 40, Appl  
Sequence 23, Appl  
Sequence 7, Appl  
Sequence 54, Appl

```
Query Match          50.9%; Score 84; DB 2; Length 25;
Best Local Similarity 63.6%; Pred. No. 4.1e-06;
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31
   : : : : |||||
Db 3 VDVKDCANHEIKKVLVPGCHGS 24

RESULT 2
US-08-478-572-169
; Sequence 169, Application US/08478572
; Patent No. 5968526
; GENERAL INFORMATION:
; APPLICANT: Garman, Richard
; APPLICANT: Greenstein, Julia
; APPLICANT: Kuo, Mei-chang
; APPLICANT: Rogers, Bruce
; APPLICANT: Franzen, Henry
; APPLICANT: Chen, Xian
; APPLICANT: Evans, Sean
; APPLICANT: Shaked, Ze'ev
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS
; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,296
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/445,307
; FILING DATE: 07 June 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAIG, ANNE I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: 017.6US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6040
; INFORMATION FOR SEQ ID NO: 169:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-478-572-169

Query Match          50.9%; Score 84; DB 2; Length 25;
Best Local Similarity 63.6%; Pred. No. 4.1e-06;
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31
   : : : : |||||
Db 3 VDVKDCANHEIKKVLVPGCHGS 24

RESULT 3
US-08-484-296-169
; Sequence 169, Application US/08484296
; Patent No. 6268491
; GENERAL INFORMATION:
; APPLICANT: Garman, Richard
; APPLICANT: Greenstein, Julia
; APPLICANT: Kuo, Mei-chang
; APPLICANT: Rogers, Bruce
; APPLICANT: Franzen, Henry
; APPLICANT: Chen, Xian
; APPLICANT: Evans, Sean
; APPLICANT: Shaked, Ze'ev
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS
; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,296
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/445,307
; FILING DATE: 07 June 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAIG, ANNE I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: 017.6US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6040
; INFORMATION FOR SEQ ID NO: 169:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-484-296-169

Query Match          50.9%; Score 84; DB 4; Length 25;
Best Local Similarity 63.6%; Pred. No. 4.1e-06;
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31
   : : : : |||||
Db 3 VDVKDCANHEIKKVLVPGCHGS 24

RESULT 4
PCT-US95-04481-31
; Sequence 31, Application PC/TUS9504481
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Pharmaceutical Peptide Formulations For Treatment of Dust M
; NUMBER OF SEQUENCES: 54
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04481
; FILING DATE:
; CLASSIFICATION:
```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/227,772
; FILING DATE: April 14, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Vanstone, Darlene A.
; REGISTRATION NUMBER: 35,279
; REFERENCE/DOCKET NUMBER: 017.5 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6040
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-04481-31

Query Match 50.9%; Score 84; DB 5; Length 25;
Best Local Similarity 63.6%; Pred. No. 4.1e-06;
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 10 IGITELGGHEIKKVLVPGCHGS 31
: : : |||||
Db 3 VDVKDCANHEIKKVLVPGCHGS 24

RESULT 5
US-08-482-142-50
; Sequence 50, Application US/08482142
; Patent No. 5820862
; GENERAL INFORMATION:
; APPLICANT: Garman, Richard
; APPLICANT: Greenstein, Julia
; APPLICANT: Kuo, Mei-chang
; APPLICANT: Rogers, Bruce
; APPLICANT: Franzen, Henry
; APPLICANT: Chen, Xian
; APPLICANT: Evans, Sean
; APPLICANT: Shaked, Ze'ev
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS
; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,142
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,307
; FILING DATE: 07 June 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAIG, ANNE I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: 017.6US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6040
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids

Query Match 50.9%; Score 84; DB 5; Length 25;
Best Local Similarity 63.6%; Pred. No. 4.1e-06;
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 10 IGITELGGHEIKKVLVPGCHGS 31
: : : |||||
Db 3 VDVKDCANHEIKKVLVPGCHGS 24

RESULT 6
US-08-482-142-170
; Sequence 170, Application US/08482142
; Patent No. 5820862
; GENERAL INFORMATION:
; APPLICANT: Garman, Richard
; APPLICANT: Greenstein, Julia
; APPLICANT: Kuo, Mei-chang
; APPLICANT: Rogers, Bruce
; APPLICANT: Franzen, Henry
; APPLICANT: Chen, Xian
; APPLICANT: Evans, Sean
; APPLICANT: Shaked, Ze'ev
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS
; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,142
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,307
; FILING DATE: 07 June 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAIG, ANNE I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: 017.6US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6040
; INFORMATION FOR SEQ ID NO: 170:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-482-142-170

Query Match 50.9%; Score 84; DB 2; Length 26;
Best Local Similarity 63.6%; Pred. No. 4.2e-06;
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 10 IGITELGGHEIKKVLVPGCHGS 31
: : : |||||
Db 3 VDVKDCANHEIKKVLVPGCHGS 24
```

```

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
US-08-482-142-50

Query Match 50.9%; Score 84; DB 2; Length 26;
Best Local Similarity 63.6%; Pred. No. 4.2e-06;
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 10 IGITELGGHEIKKVLVPGCHGS 31
: : : |||||
Db 3 VDVKDCANHEIKKVLVPGCHGS 24

RESULT 6
US-08-482-142-170
; Sequence 170, Application US/08482142
; Patent No. 5820862
; GENERAL INFORMATION:
; APPLICANT: Garman, Richard
; APPLICANT: Greenstein, Julia
; APPLICANT: Kuo, Mei-chang
; APPLICANT: Rogers, Bruce
; APPLICANT: Franzen, Henry
; APPLICANT: Chen, Xian
; APPLICANT: Evans, Sean
; APPLICANT: Shaked, Ze'ev
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS
; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,142
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,307
; FILING DATE: 07 June 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAIG, ANNE I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: 017.6US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6040
; INFORMATION FOR SEQ ID NO: 170:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-482-142-170

Query Match 50.9%; Score 84; DB 2; Length 26;
Best Local Similarity 63.6%; Pred. No. 4.2e-06;
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 10 IGITELGGHEIKKVLVPGCHGS 31
: : : |||||
Db 3 VDVKDCANHEIKKVLVPGCHGS 24
```

RESULT 7  
US-08-478-572-50  
; Sequence 50, Application US/08478572  
; Patent No. 5968526  
; GENERAL INFORMATION:  
; APPLICANT: Garman, Richard  
; APPLICANT: Greenstein, Julia  
; APPLICANT: Kuo, Mei-chang  
; APPLICANT: Rogers, Bruce  
; APPLICANT: Franzen, Henry  
; APPLICANT: Chen, Xian  
; APPLICANT: Evans, Sean  
; APPLICANT: Shaked, Ze'ev  
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
; FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
; NUMBER OF SEQUENCES: 207  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
; STREET: 610 LINCOLN STREET  
; CITY: WALTHAM  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/478,572  
; FILING DATE: 07-June-1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/445,307  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CRAIG, ANNE I.  
; REGISTRATION NUMBER: 32,976  
; REFERENCE/DOCKET NUMBER: 017.6US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 466-6000  
; TELEFAX: (617) 466-6040  
; INFORMATION FOR SEQ ID NO: 50:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 26 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: N-terminal  
; FRAGMENT TYPE: N-terminal  
US-08-478-572-50

Query Match 50.9%; Score 84; DB 2; Length 26;  
Best Local Similarity 63.6%; Pred. No. 4.2e-06;  
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
QY 10 IGITELGGHEIKKVLVPGCHGS 31  
; : : |||||  
Db 3 VDVKDCANHEIKKVLVPGCHGS 24  
; : : |||||  
RESULT 8  
US-08-478-572-170  
; Sequence 170, Application US/08478572  
; Patent No. 5968526  
; GENERAL INFORMATION:  
; APPLICANT: Garman, Richard  
; APPLICANT: Greenstein, Julia  
; APPLICANT: Kuo, Mei-chang  
; APPLICANT: Rogers, Bruce  
; APPLICANT: Franzen, Henry  
; APPLICANT: Chen, Xian  
; APPLICANT: Evans, Sean  
; APPLICANT: Shaked, Ze'ev  
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
; FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
; NUMBER OF SEQUENCES: 207  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
; STREET: 610 LINCOLN STREET  
; CITY: WALTHAM  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/478,572  
; FILING DATE: 07-June-1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/445,307  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CRAIG, ANNE I.  
; REGISTRATION NUMBER: 32,976  
; REFERENCE/DOCKET NUMBER: 017.6US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 466-6000  
; TELEFAX: (617) 466-6040  
; INFORMATION FOR SEQ ID NO: 50:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 26 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: N-terminal  
US-08-478-572-50

Query Match 50.9%; Score 84; DB 2; Length 26;  
Best Local Similarity 63.6%; Pred. No. 4.2e-06;  
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31  
; : : |||||  
Db 3 VDVKDCANHEIKKVLVPGCHGS 24  
; : : |||||

RESULT 8  
US-08-478-572-170  
; Sequence 170, Application US/08478572  
; Patent No. 5968526  
; GENERAL INFORMATION:  
; APPLICANT: Garman, Richard  
; APPLICANT: Greenstein, Julia  
; APPLICANT: Kuo, Mei-chang  
; APPLICANT: Rogers, Bruce  
; APPLICANT: Franzen, Henry  
; APPLICANT: Chen, Xian  
; APPLICANT: Evans, Sean  
; APPLICANT: Shaked, Ze'ev  
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
; FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
; NUMBER OF SEQUENCES: 207  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
; STREET: 610 LINCOLN STREET  
; CITY: WALTHAM  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02154  
; COMPUTER READABLE FORM:

; APPLICANT: Evans, Sean  
; APPLICANT: Shaked, Ze'ev  
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
; FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
; NUMBER OF SEQUENCES: 207  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
; STREET: 610 LINCOLN STREET  
; CITY: WALTHAM  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/478,572  
; FILING DATE: 07-June-1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/445,307  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CRAIG, ANNE I.  
; REGISTRATION NUMBER: 32,976  
; REFERENCE/DOCKET NUMBER: 017.6US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 466-6000  
; TELEFAX: (617) 466-6040  
; INFORMATION FOR SEQ ID NO: 170:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 26 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: internal  
US-08-478-572-170

Query Match 50.9%; Score 84; DB 2; Length 26;  
Best Local Similarity 63.6%; Pred. No. 4.2e-06;  
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31  
; : : |||||  
Db 3 VDVKDCANHEIKKVLVPGCHGS 24  
; : : |||||

RESULT 9  
US-08-484-296-50  
; Sequence 50, Application US/08484296  
; Patent No. 6268491  
; GENERAL INFORMATION:  
; APPLICANT: Garman, Richard  
; APPLICANT: Greenstein, Julia  
; APPLICANT: Kuo, Mei-chang  
; APPLICANT: Rogers, Bruce  
; APPLICANT: Franzen, Henry  
; APPLICANT: Chen, Xian  
; APPLICANT: Evans, Sean  
; APPLICANT: Shaked, Ze'ev  
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
; FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
; NUMBER OF SEQUENCES: 207  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
; STREET: 610 LINCOLN STREET  
; CITY: WALTHAM  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02154  
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
FILING DATE: US/08/484,296  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/445,307  
FILING DATE: 07 June 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: CRAIG, ANNE I.  
REGISTRATION NUMBER: 32,976  
REFERENCE/DOCKET NUMBER: 017.605  
TELEPHONE: (617) 466-6000  
TELEFAX: (617) 466-6040  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 26 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal  
US-08-484-296-50

Query Match 50.9%; Score 84; DB 4; Length 26;  
Best Local Similarity 63.6%; Pred. No. 4.2e-06;  
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 10 IGITELGGHEIKKVLVPGCHGS 31  
Db 3 VDKDCANHEIKKVLVPGCHGS 24  
: : : |||||

RESULT 10  
US-08-484-296-170  
Sequence 170, Application US/08484296  
Patent No. 6268491  
GENERAL INFORMATION:  
APPLICANT: Garman, Richard  
APPLICANT: Greenstein, Julia  
APPLICANT: Kuo, Mei-chang  
APPLICANT: Rogers, Bruce  
APPLICANT: Franzen, Henry  
APPLICANT: Chen, Xian  
APPLICANT: Evans, Sean  
APPLICANT: Shaked, Ze'ev  
TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
NUMBER OF SEQUENCES: 207  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
STREET: 610 LINCOLN STREET  
CITY: WALTHAM  
STATE: MA  
COUNTRY: USA  
ZIP: 02154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
FILING DATE: US/08/484,296  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/445,307  
FILING DATE: 07 June 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: CRAIG, ANNE I.

REGISTRATION NUMBER: 32,976  
REFERENCE/DOCKET NUMBER: 017.605  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 466-6000  
TELEFAX: (617) 466-6040  
INFORMATION FOR SEQ ID NO: 170:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 26 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-484-296-170

Query Match 50.9%; Score 84; DB 4; Length 26;  
Best Local Similarity 63.6%; Pred. No. 4.2e-06;  
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 10 IGITELGGHEIKKVLVPGCHGS 31  
Db 3 VDKDCANHEIKKVLVPGCHGS 24  
: : : |||||

RESULT 11  
US-08-482-142-171  
Sequence 171, Application US/08482142  
Patent No. 5820862  
GENERAL INFORMATION:  
APPLICANT: Garman, Richard  
APPLICANT: Greenstein, Julia  
APPLICANT: Kuo, Mei-chang  
APPLICANT: Rogers, Bruce  
APPLICANT: Franzen, Henry  
APPLICANT: Chen, Xian  
APPLICANT: Evans, Sean  
APPLICANT: Shaked, Ze'ev  
TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
NUMBER OF SEQUENCES: 207  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
STREET: 610 LINCOLN STREET  
CITY: WALTHAM  
STATE: MA  
COUNTRY: USA  
ZIP: 02154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
FILING DATE: US/08/482,142  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/445,307  
FILING DATE: 07 June 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: CRAIG, ANNE I.  
REGISTRATION NUMBER: 32,976  
REFERENCE/DOCKET NUMBER: 017.605  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 466-6000  
TELEFAX: (617) 466-6040  
INFORMATION FOR SEQ ID NO: 171:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 27 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-482-142-171

Query Match 50.9%; Score 84; DB 2; Length 27;  
Best Local Similarity 63.6%; Pred. No. 4.4e-06;  
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31  
: : : |||||  
Db 3 VDVKDCANHEIKKVLVPGCHGS 24

RESULT 12  
US-08-478-572-171  
; Sequence 171, Application US/08478572  
; Patent No. 5968526  
; GENERAL INFORMATION:  
; APPLICANT: Garman, Richard  
; APPLICANT: Greenstein, Julia  
; APPLICANT: Kuo, Mei-chang  
; APPLICANT: Rogers, Bruce  
; APPLICANT: Franzen, Henry  
; APPLICANT: Chen, Xian  
; APPLICANT: Evans, Sean  
; APPLICANT: Shaked, Ze'ev  
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
; FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
; NUMBER OF SEQUENCES: 207  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
; STREET: 610 LINCOLN STREET  
; CITY: WALTHAM  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/478,572  
; FILING DATE: 07-June-1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/445,307  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CRAIG, ANNE I.  
; REGISTRATION NUMBER: 32,976  
; REFERENCE/DOCKET NUMBER: 017.6US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 466-6000  
; TELEFAX: (617) 466-6040  
; INFORMATION FOR SEQ ID NO: 171:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 27 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: internal  
US-08-478-572-171

Query Match 50.9%; Score 84; DB 2; Length 27;  
Best Local Similarity 63.6%; Pred. No. 4.4e-06;  
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31  
: : : |||||  
Db 3 VDVKDCANHEIKKVLVPGCHGS 24

RESULT 13  
US-08-484-296-171  
; Sequence 171, Application US/08484296

Patent No. 6268491  
; GENERAL INFORMATION:  
; APPLICANT: Garman, Richard  
; APPLICANT: Greenstein, Julia  
; APPLICANT: Kuo, Mei-chang  
; APPLICANT: Rogers, Bruce  
; APPLICANT: Franzen, Henry  
; APPLICANT: Chen, Xian  
; APPLICANT: Evans, Sean  
; APPLICANT: Shaked, Ze'ev  
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
; FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
; NUMBER OF SEQUENCES: 207  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
; STREET: 610 LINCOLN STREET  
; CITY: WALTHAM  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,296  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/445,307  
; FILING DATE: 07 June 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CRAIG, ANNE I.  
; REGISTRATION NUMBER: 32,976  
; REFERENCE/DOCKET NUMBER: 017.6US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 466-6000  
; TELEFAX: (617) 466-6040  
; INFORMATION FOR SEQ ID NO: 171:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 27 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: internal  
US-08-484-296-171

Query Match 50.9%; Score 84; DB 4; Length 27;  
Best Local Similarity 63.6%; Pred. No. 4.4e-06;  
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31  
: : : |||||  
Db 3 VDVKDCANHEIKKVLVPGCHGS 24

RESULT 14  
US-08-482-142-52  
; Sequence 52, Application US/08482142  
; Patent No. 5820862  
; GENERAL INFORMATION:  
; APPLICANT: Garman, Richard  
; APPLICANT: Greenstein, Julia  
; APPLICANT: Kuo, Mei-chang  
; APPLICANT: Rogers, Bruce  
; APPLICANT: Franzen, Henry  
; APPLICANT: Chen, Xian  
; APPLICANT: Evans, Sean  
; APPLICANT: Shaked, Ze'ev  
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
; FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
; NUMBER OF SEQUENCES: 207



```
;
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482.142
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/445.307
; FILING DATE: 07 June 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAIG, ANNE I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: 017.60S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6040
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; US-08-482-142-52

Query Match 50.3%; Score 83; DB 2; Length 26;
Best Local Similarity 63.68; Pred. No. 6e-06;
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31
: : : |||||
Db 3 VDVKDSANHEIKKVLVPGCHGS 24

RESULT 15
US-08-478-572-52
; Sequence 52, Application US/08478572
; Patent No. 5968526
; GENERAL INFORMATION:
; APPLICANT: Garman, Richard
; APPLICANT: Greenstein, Julia
; APPLICANT: Kuo, Mei-chang
; APPLICANT: Rogers, Bruce
; APPLICANT: Franzen, Henry
; APPLICANT: Chen, Xian
; APPLICANT: Evans, Sean
; APPLICANT: Shaked, Ze'ev
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS
; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
```

```
;
;
; APPLICATION NUMBER: US/08/478,572
; FILING DATE: 07-June-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/445,307
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAIG, ANNE I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: 017.60S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6040
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; US-08-478-572-52

Query Match 50.3%; Score 83; DB 2; Length 26;
Best Local Similarity 63.6%; Pred. No. 6e-06;
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31
: : : |||||
Db 3 VDVKDSANHEIKKVLVPGCHGS 24

Search completed: December 4, 2002, 13:07:45
Job time : 16 secs
```



Dbb 6 IKQOIILOQVSTSVLAQAKRPKEV 29

RESULT 3  
A55224

hypothetical protein W - Rhodobacter sphaeroides (fragment)  
C:Species: Rhodobacter sphaeroides  
C:Date: 25-Aug-1995 #sequence\_revision 25-Aug-1995 #text\_change 08-Oct-1999  
C:Accession: A55224  
R: Xu, H.H.; Tabita, F.R.  
J. Bacteriol. 176, 7299-7308, 1994  
A:Title: Positive and negative regulation of sequences upstream of the form II cbb CO-2  
A:Reference number: A55224; MUID:95050314; PMID:7961502  
A:Accession: A55224  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-38 <XUA>  
A:Cross-references: GB:U12430; NID:g607900; PIDN:AAA65074.1; PID:g765112

Query Match 17.2%; Score 35; DB 2; Length 38;  
Best Local Similarity 31.6%; Pred. No. 1e+03;  
Matches 6; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 19 KYVQNTLKLATGKGVII 37  
::: | | | | | | | | | |

Dbb 9 RHREANALRTSGESGEL 27

RESULT 4  
I48113

histone H3.2 - Chinese hamster (fragment)  
C:Species: Cricetulus griseus (Chinese hamster)  
C:Date: 04-Sep-1997 #sequence\_revision 07-Nov-1997 #text\_change 21-Jul-2000  
C:Accession: I48113  
R: Artishevsky, A.; Wooden, S.; Sharma, A.; Resendez, E.  
Nature 328, 823-827, 1987  
A:Title: Cell-cycle regulatory sequences in a hamster histone promoter and their interaction  
A:Reference number: I48113; MUID:87315341; PMID:3627229  
A:Accession: I48113  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-39 <RES>  
A:Cross-references: GB:M28265; NID:g341874; PIDN:AAA42371.1; PID:g554580  
C:Genetics:  
A:Gene: H3.2  
C:Superfamily: histone H3

Query Match 17.2%; Score 35; DB 2; Length 39;  
Best Local Similarity 31.4%; Pred. No. 1e+03;  
Matches 11; Conservative 5; Mismatches 17; Indels 2; Gaps 1;

QY 6 QNTLKLATGKGGPKYVKQNTLKLATGKGVIIIGIK 40  
: | : | | | | | | | | | |

Dbb 5 KOTARKSTGGKAPR--KOLATRAAKRSAPATGGVK 37

RESULT 5  
C82475

hypothetical protein VCA0321 [imported] - Vibrio cholerae (strain N16961 serogroup O1)  
C:Species: Vibrio cholerae  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C:Accession: C82475  
R: Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.;  
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers,  
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: A82035; MUID:20406833; PMID:10952301  
A:Accession: C82475  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-36 <HEI>  
A:Cross-references: GB:AE004370; GB:AE003853; NID:g9657704; PIDN:AAF96229.1; GSPDB:GN001

A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VCA0321  
A:Map position: 2

Query Match 16.7%; Score 34; DB 2; Length 36;  
Best Local Similarity 47.1%; Pred. No. 1.3e+03;  
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 9 LKLTATGKGGPKYVKQNT 25  
: | : | | | | | | | |

Dbb 18 MHLASTRPTPKAVKQAT 34

RESULT 6  
B82151

hypothetical protein VCL830 [imported] - Vibrio cholerae (strain N16961 serogroup O1)  
C:Species: Vibrio cholerae  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C:Accession: B82151  
R: Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.;  
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers,  
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: A82035; MUID:20406833; PMID:10952301  
A:Accession: B82151  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-31 <HEI>  
A:Cross-references: GB:AE004259; GB:AE003852; NID:g9656353; PIDN:AAF94978.1; GSPDB:GN  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VCL830  
A:Map position: 1

Query Match 16.3%; Score 33; DB 2; Length 31;  
Best Local Similarity 36.4%; Pred. No. 1.4e+03;  
Matches 8; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

QY 13 TGKKGPKYVKQNTLKLATGKKG 34  
| | | | | : | | |

Dbb 8 THKKGRKEGRNVIYSTDGKDG 29

RESULT 7  
I37460

histone H3 - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 01-Nov-1996 #sequence\_revision 01-Nov-1996 #text\_change 23-Jul-1999  
C:Accession: I37460  
R: Clark, S.J.; Krieg, P.A.; Wells, J.R.  
Nucleic Acids Res. 9, 1583-1590, 1981  
A:Title: Isolation of a clone containing human histone genes.  
A:Reference number: I37460; MUID:81198981; PMID:6262733  
A:Accession: I37460  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-27 <RES>  
A:Cross-references: EMBL:V00521; NID:g32094; PIDN:CAA23780.1; PID:g32095  
C:Superfamily: histone H3

Query Match 15.8%; Score 32; DB 2; Length 27;  
Best Local Similarity 37.5%; Pred. No. 1.7e+03;  
Matches 9; Conservative 4; Mismatches 9; Indels 2; Gaps 1;

QY 6 QNTLKLATGKGGPKYVKQNTLKLAA 29  
: | : | | | | | | | |

Dbb 5 KOTARKSTGGKAPR--KOLATKAA 26

RESULT 8  
A39525

```
histone H3 - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 24-Jan-1992 #sequence_revision 24-Jan-1992 #text_change 12-Apr-1995
C:Accession: A39525
R:Mahadevan, L.C.; Willis, A.C.; Barratt, M.J.
Cell 65, 775-783, 1991
A:Title: Rapid histone H3 phosphorylation in response to growth factors, phorbol esters.
A:Reference number: A39525; MUID:91249384; PMID:2040014
A:Accession: A39525
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-35 <NAH>
C:Superfamily: histone H3
C:Keywords: Chromosomal protein; nucleosome core

Query Match 15.8%; Score 32; DB 2; Length 35;
Best Local Similarity 37.5%; Pred. NO. 2.1e+03;
Matches 9; Conservative 4; Mismatches 9; Indels 2; Gaps 1;

QY 6 QNTLKLATGKKGPKYVKQNTLKL 29
   : : : : : : : : : :
Db 4 KOTARKSTGGKAPR--KOLATKAA 25

RESULT 9
S35676
protein kinase - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 10-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 10-Nov-1995
C:Accession: S35676
R:Brunati, A.M.; James, P.; Donella-Deana, A.; Matoskova, B.; Robbins, K.C.; Pinna, L.A.
Eur. J. Biochem. 216, 323-327, 1993
A:Title: Isolation and identification of two proto-oncogene products related to c-fgr an
A:Reference number: S35675; MUID:93373942; PMID:8365414
A:Accession: S35676
A:Molecule type: protein
A:Residues: 1-21 <BRU>

Query Match 15.3%; Score 31; DB 2; Length 21;
Best Local Similarity 38.9%; Pred. NO. 1.7e+03;
Matches 7; Conservative 4; Mismatches 4; Indels 2; Gaps 1;

QY 7 NTLKLATGKKGPKYVKQNT 24
   : : : : : : : :
Db 5 NSSELTGTGTG--YIPSN 20

RESULT 10
PQ0697
hemagglutinin [imported] - rice (fragment)
C:Species: Oryza sativa (rice)
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: PQ0697
R:Komatsu, S.; Kajiura, H.; Hirano, H.
Theor. Appl. Genet. 86, 935-942, 1993
A:Title: A rice protein library; a data-file of rice proteins separated by two-dimension
A:Reference number: PQ0696
A:Accession: PQ0697
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-22 <KOM>

Query Match 15.3%; Score 31; DB 2; Length 22;
Best Local Similarity 42.1%; Pred. NO. 1.8e+03;
Matches 8; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 17 GPKYVKQNTLKLATGKKGV 35
   : : : : : : : :
Db 1 GPKFVVGSLKLGSKDTV 19

RESULT 11
S09297
```

```
cytochrome-c oxidase (EC 1.9.3.1) chain Vc - sweet potato (fragment)
C:Species: Ipomoea batatas (sweet potato)
C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 23-Feb-1997
C:Accession: S09297
R:Nakagawa, T.; Maeshima, M.; Muto, H.; Kajiura, H.; Hattori, H.; Asahi, T.
Eur. J. Biochem. 165, 303-307, 1987
A:Title: Separation, amino-terminal sequence and cell-free synthesis of the smallest
A:Reference number: S09297; MUID:87246607; PMID:2885192
A:Accession: S09297
A:Molecule type: protein
A:Residues: 1-25 <NAK>
C:Keywords: mitochondrion; oxidoreductase

Query Match 15.3%; Score 31; DB 2; Length 25;
Best Local Similarity 75.0%; Pred. NO. 2.1e+03;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 16 KGPKYVKQ 23
   : : : : :
Db 11 KGPSVVKQ 18

RESULT 12
PC2300
gaegurin 1 - Korean frog (fragment)
C:Species: Rana rugosa (Korean frog)
C>Date: 25-Feb-1995 #sequence_revision 26-May-1995 #text_change 19-May-2000
C:Accession: PC2300
R:Park, J.M.; Jung, J.E.; Lee, R.J.
Biochem. Biophys. Res. Commun. 205, 948-954, 1994
A:Title: Antimicrobial peptides from the skin of a Korean frog, Rana rugosa.
A:Reference number: PC2300; MUID:95091844; PMID:7999137
A:Accession: PC2300
A:Molecule type: protein
A:Residues: 1-33 <PAR>
C:Superfamily: ranalexin precursor; dermorphin precursor amino-terminal homology
C:Keywords: skin
F:27-33/Region: rana box motif

Query Match 15.3%; Score 31; DB 2; Length 33;
Best Local Similarity 46.2%; Pred. NO. 2.7e+03;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 15 KKGPKYVKQNTLK 27
   : : : : : : :
Db 7 KAGAKFLGKLLK 19

RESULT 13
S58601
hypothetical protein 38 - maize chloroplast
C:Species: chloroplast Zea mays (maize)
C>Date: 29-Nov-1995 #sequence_revision 19-Jan-1996 #text_change 29-Oct-1999
C:Accession: S58601
R:Maier, R.M.; Neckeremann, K.; Igloi, G.L.; Koessel, H.
J. Mol. Biol. 251, 614-628, 1995
A:Title: Complete sequence of the maize chloroplast genome: gene content, hotspots of
A:Reference number: S58531; MUID:95395841; PMID:7666415
A:Accession: S58601
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-38 <MAI>
A:Cross-references: EMBL:X86563; NID:g902200; PIDN:CAA60335.1; PID:g902270
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1995
C:Genetics:
A:Genome: chloroplast
C:Keywords: chloroplast

Query Match 15.0%; Score 30.5; DB 2; Length 38;
Best Local Similarity 29.6%; Pred. NO. 3.5e+03;
Matches 8; Conservative 7; Mismatches 9; Indels 3; Gaps 1;

QY 10 KLATG---KKGPKYVKQNTLKLATGKK 33
```

```
Db      4  RLSTSTWRKKPKKHMASNWFTRSTQER 30
      :|:|  || ||:|  |  :|:|
RESULT 14
S19618
globin - polychaete (Eudistyllia vancouveri) (fragment)
N:Alternate names: chlorocruorin
C:Species: Eudistyllia vancouveri
C:Date: 19-Mar-1997 #sequence_revision 10-Oct-1997 #text_change 05-Dec-1998
C:Accession: S19618
R:Qabar, A.N.; Stern, M.S.; Walz, D.A.; Chiu, J.T.; Timkovich, R.; Wall, J.S.; Kapp, O.H.
J. Mol. Biol. 222, 1109-1129, 1991
A:Title: Hierarchy of globin complexes. The quaternary structure of the extracellular ch
A:Reference number: S19532; MUID:92106333; PMID:1762147
A:Accession: S19618
A:Molecule type: protein
A:Residues: 1-20 <QAB>
A:Experimental source: plume
C:Complex: dodecamers, each consisting of a trimer of tetramers of globin chains; dodeca
C:Keywords: calcium; dodecamer; heme; homotetramer; oxygen carrier
      Query Match      14.8%; Score 30; DB 2; Length 20;
      Best Local Similarity 75.0%; Pred. No. 2.2e+03;
      Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 12  ATGKKGPK 19
      | |||| |
Db 12  ANGKKGDK 19
RESULT 15
B49732
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) 38.5K chain - potato mitochondrion (fragme
N:Alternate names: complex I dehydrogenase 38.5K chain; NADH-ubiquinone oxidoreductase 3
C:Species: mitochondrion Solanum tuberosum (potato)
C:Date: 12-May-1994 #sequence_revision 12-May-1994 #text_change 03-Jun-2002
C:Accession: B49732
R:Herz, U.; Schroeder, W.; Liddell, A.; Leaver, C.J.; Brennicke, A.; Grohmann, L.
J. Biol. Chem. 269, 2263-2269, 1994
A:Title: Purification of the NADH:ubiquinone oxidoreductase (complex I) of the respirato
A:Reference number: A49732; MUID:94124587; PMID:8294484
A:Accession: B49732
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-22 <HER>
C:Genetics:
A:Genome: mitochondrion
C:Keywords: electron transfer; hydrogen ion transport; mitochondrial inner membrane; mit
      Query Match      14.8%; Score 30; DB 2; Length 22;
      Best Local Similarity 58.3%; Pred. No. 2.4e+03;
      Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 11  LATGKKGPYVK 22
      |||| | |
Db 4   LATGGAGPLIX 15
```

Search completed: December 4, 2002, 13:12:46  
Job time : 19 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 13:07:51 ; Search time 11 seconds

(without alignments)  
150.823 Million cell updates/sec

Title: US-09-362-731A-4

Perfect score: 203

Sequence: 1 PKYVKQNTLKLATCKGPKYVKQNTLKLATCKGKGVIGIK 40

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 2868

Minimum DB seq length: 0

Maximum DB seq length: 40

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	17.7	40	1 H32_TETAM	P17705 tetrahymena
2	33.5	16.5	23	1 A044_LITRA	P82400 litoria ran
3	32	15.8	40	1 H32_TETBO	P17319 tetrahymena
4	31.5	15.5	30	1 PG3_PACGO	P82416 pachycondyl
5	31	15.3	26	1 M031_LITGE	P82069 litoria gen
6	31	15.3	33	1 GAEL_RANRU	P80395 rana rugosa
7	31	15.3	33	1 RUGB_RANRU	P80955 rana rugosa
8	30.5	15.0	30	1 PG5_PACGO	P82418 pachycondyl
9	30	14.8	22	1 NEM_SOLTU	P80265 solanum tub
10	30	14.8	30	1 COXC_SOLTU	P80500 solanum tub
11	30	14.8	37	1 PIP7_BOVIN	P21671 bos taurus
12	29	14.3	40	1 Y41_BPT7	P03782 bacterioph
13	28.5	14.0	34	1 RR2_OCHNE	Q40606 ochrosphaer
14	28	13.8	24	1 CT31_LITCI	P81851 litoria cit
15	28	13.8	31	1 PSAM_CHLVU	P56314 chlorella v
16	28	13.8	33	1 RL21_XENLA	P49628 xenopus lae
17	28	13.8	36	1 Y609_BORBU	O51554 borrelia bu
18	28	13.8	40	1 H2B3_ECHES	P13283 echinus esc
19	27.5	13.5	32	1 ATP7_SPIOL	P80088 spinacia ol
20	27	13.3	23	1 CH41_LITCE	P56242 litoria cae
21	27	13.3	29	1 ITH3_BOVIN	P56652 bos taurus
22	27	13.3	32	1 LEC_DOLAX	P02875 dolichos ax
23	27	13.3	36	1 AMPL_PIG	P28839 sus scrofa
24	27	13.3	38	1 Y520_BORBU	O51470 borrelia bu
25	26.5	13.1	21	1 NPH_RAT	P58522 rattus norv
26	26.5	13.1	40	1 Y665_HAEIN	P44283 haemophilus
27	26	12.8	23	1 CR43_LITCE	P56244 litoria cae
28	26	12.8	28	1 FIBA_CANFA	P02673 canis famil
29	26	12.8	29	1 PG4_PACGO	P82417 pachycondyl
30	26	12.8	30	1 PG2_PACGO	P82415 pachycondyl
31	26	12.8	35	1 THPA_THADA	P21381 thaumatococ
32	26	12.8	35	1 VORB_METTM	P80908 methanobact
33	26	12.8	36	1 ANFV_ANGJA	P22642 anguilla ja

#### ALIGNMENTS

##### RESULT 1

H32_TETAM	STANDARD;	PRT;	40 AA.
ID	H32_TETAM	STANDARD;	PRT;
AC	P17705;		
DT	01-AUG-1990 (Rel. 15, Created)		
DT	01-AUG-1990 (Rel. 15, Last sequence update)		
DT	15-JUL-1999 (Rel. 38, Last annotation update)		
DE	Histone H3.2 (Fragment).		
OS	Tetrahymena americanus,		
OS	Tetrahymena australis,		
OS	Tetrahymena capricornis,		
OS	Tetrahymena caudata,		
OS	Tetrahymena elliotti,		
OS	Tetrahymena fergusonii,		
OS	Tetrahymena hyperangularis,		
OS	Tetrahymena leucophrys,		
OS	Tetrahymena malaccensis,		
OS	Tetrahymena mimbres,		
OS	Tetrahymena nanneyi,		
OS	Tetrahymena nipissingi,		
OS	Tetrahymena paravorax,		
OS	Tetrahymena patula,		
OS	Tetrahymena pigmentosa,		
OS	Tetrahymena rostrata,		
OS	Tetrahymena sonneborni, and		
OS	Tetrahymena tropicalis.		
OC	Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;		
OC	Tetrahymenina; Tetrahymena.		
OX	NCBI_TaxID=5891, 5892, 5895, 5896, 5897, 5898, 5899, 5900, 5901, 5902,		
OX	5903, 5922, 5905, 5906, 5907, 5909, 5910, 5912;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=90221813; PubMed=2129549;		
RA	Brunk C.F., Sadler L.A.;		
RT	"Characterization of the promoter region of Tetrahymena genes.";		
RL	Nucleic Acids Res. 18:323-329(1990).		
RL	[2]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=90219078; PubMed=2129541;		
RA	Brunk C.F., Kahn R.W., Sadler L.A.;		
RT	"Phylogenetic relationships among Tetrahymena species determined		
RT	using the polymerase chain reaction.";		
RL	J. Mol. Evol. 30:290-297(1990).		
CC	-!- FUNCTION: HISTONE H3, ALONG WITH HISTONE H4, PLAYS A CENTRAL ROLE		
CC	IN NUCLEOSOME FORMATION.		
CC	-!- SUBUNIT: THE NUCLEOSOME IS AN OCTAMER CONTAINING TWO MOLECULES OF		
CC	H2A, H2B, H3, AND H4; WHICH WRAP APPROXIMATIVELY 146 BP OF DNA.		
CC	-!- SIMILARITY: BELONGS TO THE HISTONE H3 FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		

34	26	12.8	38	1	SCK1_ORTSC	P55896 orthochirus
35	26	12.8	40	1	ALF_CANAL	Q9urb4 candida alb
36	25.5	12.6	22	1	CR32_LITCE	P56239 litoria cae
37	25.5	12.6	30	1	PCGI_PACGO	P82414 pachycondyl
38	25	12.3	15	1	ASPI_LACSN	P82648 lactobacill
39	25	12.3	15	1	MALT_BACTQ	P80072 bacillus th
40	25	12.3	23	1	CYSP_TRIFO	P33403 tritrichomo
41	25	12.3	29	1	PSAK_SPIOL	P14627 spinacia ol
42	25	12.3	30	1	AMPT_BACST	P00728 bacillus st
43	25	12.3	30	1	VAA1_EQUAR	Q04236 equisetum a
44	25	12.3	30	1	VAA2_EQUAR	Q04238 equisetum a
45	25	12.3	37	1	Y762_BORBU	O51703 borrelia bu



Query Match 15.3%; Score 31; DB 1; Length 26;  
 Best Local Similarity 30.0%; Pred. No. 6.4e+02;  
 Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 21 VKQNTLKLATGKGGVIGIK 40  
 :|: | | :||: ||:  
 Db 6 IKEKLESLESLAKGVISGIQ 25

RESULT 6

ID	GAEI_RANRU	STANDARD;	PRT;	33 AA.
AC	P80395;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Gaegurin-1.			
OS	Rana rugosa (Wrinkled frog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.			
OX	NCBI_TaxID=8410;			
RN	[1]			
RP	SEQUENCE.			
RC	TISSUE=Skin;			
RX	MEDLINE=95091844; PubMed=7999137;			
RA	Park J.M., Jung J.-E., Lee B.-J.;			
RT	"Antimicrobial peptides from the skin of a Korean frog, Rana			
RL	rugosa.";			
RL	Biochem. Biophys. Res. Commun. 205:948-954(1994).			
CC	-1- FUNCTION: HAS A NON-HEMOLYTIC ACTIVITY. HAS A BROAD SPECTRUM OF			
CC	ACTIVITY AGAINST BOTH GRAM-POSITIVE AND GRAM-NEGATIVE BACTERIA,			
CC	FUNGUS AND PROTOZOA.			
CC	-1- SUBUNIT: MONOMER.			
CC	-1- SUBCELLULAR LOCATION: Secreted.			
CC	-1- TISSUE SPECIFICITY: SKIN.			
CC	-1- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN			
CC	FAMILY.			
KW	Amphibian skin; Antibiotic.			
FT	DISULFID 27 33			
SQ	SEQUENCE 33 AA; 3462 MW; 51E21E3B7B0FF536 CRC64;			
	BY SIMILARITY.			
ID	RUGB_RANRU	STANDARD;	PRT;	33 AA.
AC	P80935;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Rugosin B.			
OS	Rana rugosa (Wrinkled frog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.			
OX	NCBI_TaxID=8410;			
RN	[1]			
RP	SEQUENCE.			
RC	TISSUE=Skin;			
RX	MEDLINE=95336450; PubMed=7612013;			
RA	Suzuki S., Ohe Y., Kagegawa T., Tatamoto K.;			
RT	"Isolation and characterization of novel antimicrobial peptides,			
RT	rugosins A, B and C, from the skin of the frog, Rana rugosa.";			
RL	Biochem. Biophys. Res. Commun. 212:249-254(1995).			
CC	-1- FUNCTION: SHOWS ANTIBACTERIAL ACTIVITY AGAINST BOTH GRAM-NEGATIVE			
CC	AND GRAM-POSITIVE BACTERIA.			
CC	-1- SUBCELLULAR LOCATION: Secreted.			



```

CC -!- TISSUE SPECIFICITY: SKIN.
CC -!- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GABGURIN/RUGOSIN
CC FAMILY.
KW Antibiotic: Amphibian skin.
FT DISULFID 27 33
SQ SEQUENCE 33 AA; 3516 MW; 51FFC8E1641FF536 CRC64;

Query Match 15.3%; Score 31; DB 1; Length 33;
Best Local Similarity 46.2%; Pred. No. 8.1e+02;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 15 KGPVKYVKONTLK 27
| | | | | : | | | |
Db 7 KAGARFLGNLKLK 19

RESULT 8
PCG5_PACGO STANDARD; PRT; 30 AA.
AC PC82418;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ponericin G5.
OS Pachycondylia goeldii (Ponerine ant.).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita;
OC Aculeata; Formicidae; Ponerinae; Pachycondylia.
OX NCBI_TaxID=118888;
RN [1]
RP SEQUENCE, AND FUNCTION.
RC TISSUE=Venom;
RX MEDLINP=21264562; PubMed=11279030;
RA Orivel J., Redeker V., Le Caer J.-P., Krier F., Revol-Junelles A.-M.,
RA Longeon A., Chafotte A., Dejean A., Rossier J.;
RT "Ponericins, new antibacterial and insecticidal peptides from the
RT venom of the ant Pachycondylia goeldii.";
RL J. Biol. Chem. 276:17823-17829(2001).
CC -!- FUNCTION: HAS ANTIBACTERIAL ACTIVITY AND NON-HEMOLYTIC
CC ACTIVITIES.
CC -!- MASS SPECTROMETRY: MW=3108.11; METHOD=MALDI.
KW Antibiotic; Insect immunity.
SQ SEQUENCE 30 AA; 3109 MW; D5F19CD3041688C3 CRC64;

Query Match 15.0%; Score 30.5; DB 1; Length 30;
Best Local Similarity 47.1%; Pred. No. 8.5e+02;
Matches 8; Conservative 3; Mismatches 3; Indels 3; Gaps 1;

QY 9 LKATG---KGPVKYVK 22
| | | | | : | | | |
Db 6 VKIAGGLKKGPGILK 22

RESULT 9
NUEM_SOLTU STANDARD; PRT; 22 AA.
AC P80265;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE NADH-ubiquinone oxidoreductase 38.5 kDa subunit (EC 1.6.5.3)
DE (EC 1.6.99.3) (Complex I-38.5KD) (CI-38.5KD) (Fragment).
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE.
RC STRAIN=cv. Bintje; TISSUE=Tuber;
RX MEDLINE=94124587; PubMed=8294484;
RA Hertz U., Schroeder W., Liddell A., Leaver C.J., Brennicke A.,
RA Grohmann L.;

```

```

RT "Purification of the NADH:ubiquinone oxidoreductase (complex I) of
RT the respiratory chain from the inner mitochondrial membrane of
RT Solanum tuberosum.";
RL J. Biol. Chem. 269:2263-2269(1994).
CC -!- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
CC TO BE UBIQUINONE.
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -!- CATALYTIC ACTIVITY: NADH + acceptor = NAD(+) + reduced acceptor.
CC -!- COFACTOR: FAD; CONTAINS ONE NONCOVALENTLY BOUND FAD PER
CC POLYPEPTIDE CHAIN (BY SIMILARITY).
CC -!- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 30 DIFFERENT SUBUNITS.
CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -!- SIMILARITY: BELONGS TO THE COMPLEX I 40 kDa SUBUNIT FAMILY.
KW Oxidoreductase; NAD; Ubiquinone; Flavoprotein; FAD; Mitochondrion.
FT NON_TER 22 22
SQ SEQUENCE 22 AA; 1982 MW; EDE66F57CA7B6DCF CRC64;

Query Match 14.8%; Score 30; DB 1; Length 22;
Best Local Similarity 58.3%; Pred. No. 7.2e+02;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 11 LATGKKGPKYVK 22
| | | | | : | |
Db 4 LATGGAGPLIXK 15

RESULT 10
COXC_SOLTU STANDARD; PRT; 30 AA.
ID COXC_SOLTU STANDARD; PRT; 30 AA.
AC P80500;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Cytochrome c oxidase polypeptide Vc (EC 1.9.3.1) (Fragment).
GN COXVC.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE.
RC TISSUE=Tuber;
RX MEDLINE=97077345; PubMed=8919912;
RA Jansch L., Kruft V., Schmitz U.K., Braun H.P.;
RT "New insights into the composition, molecular mass and stoichiometry
RT of the protein complexes of plant mitochondria.";
RL Plant J. 9:357-368(1996).
CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC MITOCHONDRIAL ELECTRON TRANSPORT.
CC -!- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -!- SUBCELLULAR LOCATION: Mitochondrial inner membrane.
CC -!- SIMILARITY: SOME TO THE SMALLER CYTOCHROME C OXIDASE SUBUNITS
CC FROM VARIOUS EUKARYOTES. IT MAY CORRESPOND TO YEAST SUBUNIT VIIA
CC AND MAMMALIAN SUBUNIT VIII.
KW Oxidoreductase; Inner membrane; Transmembrane; Mitochondrion.
FT DOMAIN 1 18 SIGNAL FOR IMPORT (POTENTIAL).
FT TRANSMEM 19 >30 POTENTIAL.
FT NON_TER 30 30
SQ SEQUENCE 30 AA; 3071 MW; DFD3A49141C9D92F CRC64;

Query Match 14.8%; Score 30; DB 1; Length 30;
Best Local Similarity 37.5%; Pred. No. 9.8e+02;
Matches 6; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 16 KGPVKYVKQNTLKLATG 31
| | | | | : | | | |
Db 11 KGPVVKELVIXXLG 26

```

```
RESULT 11
PIPT7_BOVIN STANDARD; PRT; 37 AA.
AC P21671;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase delta 2
DE (EC 3.1.4.11) (PLC-delta-2) (Phospholipase C-delta-2) (PLC-85)
DE (Fragments)
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX MEDLINE=89325315; PubMed=2753038;
RA Mel drum E., Katan M., Parker P.;
RT "A novel inositol-phospholipid-specific phospholipase C. Rapid
RT purification and characterization.";
RL Eur. J. Biochem. 182:673-677(1989).
CC -!- FUNCTION: THE PRODUCTION OF THE SECOND MESSENGER MOLECULES
CC DIACYLGLYCEROL (DAG) AND INOSITOL 1,4,5-TRISPHOSPHATE (IP3) IS
CC MEDIATED BY ACTIVATED PHOSPHATIDYLINOSITOL-SPECIFIC PHOSPHOLIPASE
CC C ENZYMES.
CC -!- CATALYTIC ACTIVITY: 1-phosphatidyl-1D-myo-inositol 4,5-
CC bisphosphate + H(2)O = D-myo-inositol 1,4,5-trisphosphate +
CC diacylglycerol.
CC -!- COFACTOR: Calcium.
CC -!- MISCELLANEOUS: THERE ARE AT LEAST SIX FORMS OF PLC ENZYMES.
DR PIR: S04944; S04944.
DR InterPro: IPR000909; PI_PLC_Xdom.
DR InterPro: IPR001711; PI_PLC_Y.
DR PROSITE: PS50007; PIPLC_X_DOMAIN; PARTIAL.
DR PROSITE: PS50008; PIPLC_Y_DOMAIN; 1.
KW Hydrolase; Lipid degradation; Transducer; Calcium-binding.
FT NON_TER 1
FT NON_CONS 26
FT NON_TER 37
FT NON_TER 37
SQ SEQUENCE 37 AA; 4154 MW; 061DAD0D3DBB9106 CRC64;

Query Match 14.88; Score 30; DB 1; Length 37;
Best Local Similarity 25.08; Pred. No. 1.2e+03;
Matches 7; Conservative 9; Mismatches 8; Indels 4; Gaps 1;

QY 17 GPKYVKONTLKA---TGKKGVIIGIK 40
DB 6 GDFVQHNAXQLSRVPSGLRVEIFGVR 33

RESULT 12
Y41_BPT7 STANDARD; PRT; 40 AA.
AC P03782;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Hypothetical gene 4.1 protein.
DE Hypothetical gene 4.1 protein.
GN 4.1.
OS Bacteriophage T7.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC T7-like viruses.
OX NCBI_TaxID=10760;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83241725; PubMed=6864790;
RA Dunn J.J., Studier F.W.;
RT "Complete nucleotide sequence of bacteriophage T7 DNA and the
RT locations of T7 genetic elements.";
RL J. Mol. Biol. 166:477-535(1983).
CC -!- This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X99078; CAA67534.1; -
CC InterPro: IPR001865; Ribosomal_S2.
CC Pfam: PF00318; Ribosomal_S2; 1.
CC PROSITE: PS00962; RIBOSOMAL_S2_1; PARTIAL.
CC PROSITE: PS00963; RIBOSOMAL_S2_2; PARTIAL.
CC KW Ribosomal protein; Chloroplast.
FT NON_TER 1
SQ SEQUENCE 34 AA; 3640 MW; 2039BA0FB5710655 CRC64;

Query Match 14.08; Score 28.5; DB 1; Length 34;
Best Local Similarity 40.08; Pred. No. 1.7e+03;
Matches 8; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 2 KYVKONTLKLATG-KKGPKY 20
DB 15 KSIKLIVSKLSTGIQOGQSY 34

RESULT 14
CT31_LITCI STANDARD; PRT; 24 AA.
AC P81851; P81853;
DT 30-MAY-2000 (Rel. 39, Created)
```

30-MAY-2000 (Rel. 39, Last sequence update)  
16-OCT-2001 (Rel. 40, Last annotation update)  
Citropin 3.1.2 [Contains: Citropin 3.1.1; Citropin 3.1.1].  
Litoria citropa (Australian blue mountains tree frog).  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;  
Pelodyadinae; Litoria.  
NCBI\_TaxID=94770;  
[1]  
SEQUENCE.  
TISSUE=Skin;  
MEDLINE=99435977; PubMed=10504394;  
Wegener K.L., Wabnitz P.A., Carver J.A., Bowie J.H., Chia B.C.S.,  
Wallace J.C., Tyler M.J.;  
"Host defence peptides from the skin glands of the Australian blue  
mountains tree-frog Litoria citropa. Solution structure of the  
antibacterial peptide citropin 1.1.";  
Eur. J. Biochem. 265:627-637(1999).  
CC -!- TISSUE SPECIFICITY: DORSAL AND SUBMENTAL SKIN GLANDS.  
FT PEPTIDE 1 24 CITROPIN 3.1.2.  
FT PEPTIDE 1 23 CITROPIN 3.1.1.  
FT PEPTIDE 1 22 CITROPIN 3.1.  
SQ SEQUENCE 24 AA; 2614 MW; C9001E295BD0E15D CRC64;  
  
Query Match 13.8%; Score 28; DB 1; Length 24;  
Best Local Similarity 50.0%; Pred. No. 1.4e+03;  
Matches 10; Conservative 2; Mismatches 6; Indels 2; Gaps 1;  
  
QY 21 VKQNTLKLATGKKGVIIGIK 40  
I : || || || || ||  
DB 5 VIKKKLKLATG--GVIEGIQ 22  
  
RESULT 15  
PSAM\_CHLVU  
ID PSAM\_CHLVU STANDARD; PRT; 31 AA.  
AC P56314;  
15-JUL-1998 (Rel. 36, Created)  
15-JUL-1998 (Rel. 36, Last sequence update)  
16-OCT-2001 (Rel. 40, Last annotation update)  
DE Photosystem I reaction centre subunit XII (PSI-M).  
GN PSAM.  
OS Chlorella vulgaris.  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;  
OC Chlorellaceae; Chlorella.  
OX NCBI\_TaxID=3077;  
[1]  
SEQUENCE FROM N.A.  
RC STRAIN=IAM C-27 / Tamiya;  
RX MEDLINE=97303241; PubMed=9159184;  
RA Wakasugi T., Nagai T., Kapoor M., Sugita M., Ito M., Ito S.,  
RA Tsudzuki J., Nakashima K., Tsudzuki T., Suzuki Y., Hamada A., Ohta T.,  
RA Inamura A., Yoshinaga K., Sugiyura M.;  
RT "Complete nucleotide sequence of the chloroplast genome from the  
green alga Chlorella vulgaris: the existence of genes possibly  
involved in chloroplast division.";  
RT Proc. Natl. Acad. Sci. U.S.A. 94:5967-5972(1997).  
CC -!- SIMILARITY: BELONGS TO THE PSAM FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AB001684; BAA57938.1; -  
KW Photosystem I; Photosynthesis; Chloroplast.  
SQ SEQUENCE 31 AA; 3310 MW; 184858F3D8BD6873 CRC64;  
  
Query Match 13.8%; Score 28; DB 1; Length 31;

Best Local Similarity 50.0%; Pred. No. 1.8e+03;  
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
  
QY 26 LKLATGKKGVIIGI 39  
I ||||| I : I :  
DB 14 LALATGIFAVRLGV 27  
  
Search completed: December 4, 2002, 13:11:45  
Job time : 12 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: December 4, 2002, 13:08:11 ; Search time 28 Seconds  
(without alignments)  
294.353 Million cell updates/sec

Title: US-09-362-731A-4  
Perfect score: 203  
Sequence: 1 PKYVKQNTLKLATGKGPYVKQNTLKLATGKGVIGIK 40

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 24116

Minimum DB seq length: 0  
Maximum DB seq length: 40

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : SPTREMBL.21.\*
- 1: sp\_archaea.\*
  - 2: sp\_bacteria.\*
  - 3: sp\_fungi.\*
  - 4: sp\_human.\*
  - 5: sp\_invertebrate.\*
  - 6: sp\_mammal.\*
  - 7: sp\_mhc.\*
  - 8: sp\_organelle.\*
  - 9: sp\_phage.\*
  - 10: sp\_plant.\*
  - 11: sp\_rodent.\*
  - 12: sp\_virus.\*
  - 13: sp\_vertebrate.\*
  - 14: sp\_unclassified.\*
  - 15: sp\_rvirus.\*
  - 16: sp\_bacteriap.\*
  - 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	19.7	26	4 Q9UN14	Q9un14 homo sapien
2	38.5	19.0	37	16 Q9JTY0	Q9jty0 neisseria m
3	35	17.2	37	3 Q8TF7	O8tff7 schizosacch
4	35	17.2	38	2 Q3057	Q53057 rhodobacter
5	35	17.2	38	10 Q24165	Q24165 nicotiana t
6	35	17.2	39	11 Q92167	Q92167 cricetus
7	34.5	17.0	30	12 Q96630	Q96630 bovine aden
8	34.5	17.0	34	2 Q56234	Q56234 thermus aqu
9	34	16.7	33	5 Q9NJN4	Q9njn4 neospora ca
10	34	16.7	36	16 Q9KML6	Q9kml6 vibrio chol
11	33.5	16.5	39	11 Q9R137	Q9r137 mus musculu
12	33	16.3	28	2 Q87604	Q87604 acetobacter
13	33	16.3	30	2 Q9RQ54	Q9rq54 buchnera ap
14	33	16.3	31	16 Q9KRL7	Q9krl7 vibrio chol
15	33	16.3	37	2 Q9ZEL1	Q9zell helicobacte
16	33	16.3	40	2 Q48267	Q48267 haemophilus

17	32.5	16.0	34	16	O8VKI6	O8vki6 mycobacteri
18	32	15.8	22	2	Q9F6K5	Q9f6k5 borrelia af
19	32	15.8	27	4	Q16776	Q16776 homo sapien
20	32	15.8	30	2	P83001	P83001 pseudomonas
21	32	15.8	36	5	O61187	O61187 colpoda cuc
22	32	15.8	36	5	O62593	O62593 obertrunia
23	32	15.8	36	5	O62612	O62612 prorodon te
24	31.5	15.5	24	11	O884N1	O8r4n1 mus musculu
25	31	15.3	22	4	Q13373	Q13373 homo sapien
26	31	15.3	24	2	Q9R573	Q9r573 nitrosomona
27	31	15.3	24	2	O8RL36	O8rl36 burkholderi
28	31	15.3	26	4	O96HQ8	O96hq8 homo sapien
29	31	15.3	32	4	Q96RK7	Q96rk7 homo sapien
30	31	15.3	34	3	Q12681	Q12681 saccharomyc
31	31	15.3	37	2	Q9R5S5	Q9r5s5 chlamydia t
32	31	15.3	37	5	O97124	O97124 toxoplasma
33	31	15.3	40	4	Q16812	Q16812 homo sapien
34	31	15.3	40	10	O9LHY7	O9lhy7 oryza sativ
35	30.5	15.0	31	4	Q9UDE5	Q9ude5 homo sapien
36	30.5	15.0	37	2	O52909	O52909 campylobact
37	30.5	15.0	38	8	Q33303	Q33303 zea mays (m
38	30	14.8	29	2	Q9R4B5	Q9r4b5 porphyromon
39	30	14.8	29	10	Q9S8Z6	Q9s8z6 amaranthus
40	30	14.8	30	10	Q9S8Z7	Q9s8z7 amaranthus
41	30	14.8	31	6	Q9GKI8	Q9gki8 sus scrofa
42	30	14.8	31	9	O8W688	O8w688 bacterioph
43	30	14.8	31	16	O51172	O51172 borrelia bu
44	30	14.8	32	4	Q9BYF3	Q9byf3 homo sapien
45	30	14.8	37	12	Q91AW9	Q91aw9 hepatitis c

ALIGNMENTS

RESULT 1

Q9UN14 PRELIMINARY; PRT; 26 AA.  
ID Q9UN14  
AC Q9UN14;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Phosphodiesterase I/nucleotide pyrophosphatase 3 (EC 3.1.4.1)  
DE (Fragment).  
GN PDNP3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99453721; PubMed=10524196;  
RA Andoh K., Piao J.H., Terashima K., Nakamura H., Sano K.;  
RT "Genomic structure and promoter analysis of the ecto-phosphodiesterase  
RT I gene (PDNP3) expressed in glial cells.";  
RL Biochim. Biophys. Acta 1446:213-224(1999).  
DR EMBL; AF119714; AAD46160.1; -;  
KW Hydrolase.  
FT NON\_TER 26 26  
SQ SEQUENCE 26 AA; 2953 MW; B96AE55BDAF75E48 CRC64;

Query Match 19.7%; Score 40; DB 4; Length 26;  
Best Local Similarity 50.0%; Pred. No. 2.2e+02;  
Matches 11; Conservative 4; Mismatches 3; Indels 4; Gaps 1;

QY 6 QNTLKLATGKGPYVKQNTLK 27  
:|||||:|||||  
Db 2 ESTLTATEQP----VKKNTLK 19

RESULT 2

Q9JTY0 PRELIMINARY; PRT; 37 AA.  
ID Q9JTY0  
AC Q9JTY0;

```

DT 01-OCT-2000 (TReMBLrel. 15, Created)
DE 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Hypothetical protein NMA1583.
GN NMA1583.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z2491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=2022556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moulé S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis Z2491.";
RL Nature 404:502-506(2000).
DR EMBL; AL162756; CAB84810.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 37 AA; 4266 MW; 598916F419616958 CRC64;

Query Match 19.0%; Score 38.5; DB 16; Length 37;
Best Local Similarity 36.4%; Pred. No. 5.1e+02;
Matches 8; Conservative 5; Mismatches 2; Indels 7; Gaps 1;

QY 13 TGKKGPKY-----VKQNTLK 27
      :|| ||| :|||
Db 7 SGKPKYFQSDRWIKKNSIK 28

RESULT 3
ID Q8TF77 PRELIMINARY; PRT; 37 AA.
AC Q8TF77.
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Putative enolase, duplicated in S. pombe, S. pombe paralog enol
DE (Fragment).
GN SPAP8B6.07C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972H-;
RA Wood V., Rajandream M.A., Barrell B.G., Harris D.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL691490; CAD27913.1; -.
FT NON_TER 1
SQ SEQUENCE 37 AA; 4146 MW; 4546F2C2C4B1918B8 CRC64;

Query Match 17.2%; Score 35; DB 3; Length 37;
Best Local Similarity 46.2%; Pred. No. 1.4e+03;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 14 GKKGPKYVKQNTL 26
      | : ||| |||
Db 25 GTRAADYIKSNTL 37

RESULT 4
Q53057
ID Q53057 PRELIMINARY; PRT; 38 AA.
AC Q53057.
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)

```

```

DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Hypothetical 4.4 kDa protein (Fragment).
OS Rhodobacter sphaeroides (Rhodospirillum rubrum sphaeroides).
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Rhodobacter.
OX NCBI_TaxID=1063;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HR;
RX MEDLINE=95050314; PubMed=7961502;
RA Xu H.H., Tabita F.R.;
RT "Positive and negative regulation of sequences upstream of the form II
RT cbb CO2 fixation operon of Rhodobacter sphaeroides.";
RL J. Bacteriol. 176:7299-7308(1994).
DR EMBL; U12430; AAA65074.1; -.
KW Hypothetical protein.
FT NON_TER 38
SQ SEQUENCE 38 AA; 4403 MW; 9A0D811A9A434A97 CRC64;

Query Match 17.2%; Score 35; DB 2; Length 38;
Best Local Similarity 31.6%; Pred. No. 1.5e+03;
Matches 6; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 19 KVKQNTLKLATGKKGVII 37
      :: : ||| :|| :| :|
Db 9 RHREANALRLTSGESGEIL 27

RESULT 5
ID Q24165 PRELIMINARY; PRT; 38 AA.
AC Q24165.
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Histone H3 (Fragment).
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRIGHT YELLOW 2;
RX MEDLINE=98292547; PubMed=9628927;
RA Reichheld J.P., Gigot C., Chaubet-Gigot N.;
RT "Multilevel regulation of histone gene expression during the cell
RT cycle in plant cells.";
RL Nucleic Acids Res. 26:3255-3262(1998).
DR EMBL; Y14195; CAA74582.1; -.
DR InterPro; IPR000164; Histone_H3.
DR PROSITE; PS00322; HISTONE_H3_1; UNKNOWNW_1.
FT NON_TER 38
SQ SEQUENCE 38 AA; 3939 MW; E1600FC8D9ED8D3 CRC64;

Query Match 17.2%; Score 35; DB 10; Length 38;
Best Local Similarity 31.4%; Pred. No. 1.5e+03;
Matches 11; Conservative 5; Mismatches 17; Indels 2; Gaps 1;

QY 6 QNTLKLATGKKGPKYVKQNTLKLATGKKGVIIGIK 40
      : | : ||| :| :| :| :| :| :| :| :|
Db 5 KQIARKSTGKAPF--KQLATKAARKSAPATGGVK 37

RESULT 6
ID Q9Z1G7 PRELIMINARY; PRT; 39 AA.
AC Q9Z1G7.
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE H3.2 protein (Fragment).
GN H3.2.

```

```

OS Cricetulus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata: Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_TaxID=10031;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=87315341; PubMed=3627229;
RX Artishevsky A., Wooden S., Sharma A., Resendez E.Jr., Lee A.S.;
RT "Cell-cycle regulatory sequences in a hamster histone promoter and
RT their interactions with cellular factors.";
RL Nature 328:823-827(1987).
DR EMBL; M28265; AAA42371.1; -.
DR InterPro; IPR001164; Histone_H3.
DR PRINTS; PR00622; HISTONEH3.
DR PROSITE; PS00322; HISTONE_H3_1; UNKNOWN_1.
FT NON_TER 39
SQ SEQUENCE 39 AA; 4036 MW; DA31600FC8D9EDD8 CRC64;

Query Match 17.2%; Score 35; DB 11; Length 39;
Best Local Similarity 31.4%; Pred. No. 1.5e+03;
Matches 11; Conservative 5; Mismatches 17; Indels 2; Gaps 1;

Qy 6 QNTLKATGKGGKPKYKQNTLKATGKGGKGVIIIGK 40
: | : || | | : || | | | | | | |
Db 5 KQTARKSTGGKAPR--KQLATKAARKSAPATCGVK 37
: | : || | | : || | | | | | | |

RESULT 7
Q96630
ID Q96630 PRELIMINARY; PRT; 30 AA.
AC Q96630;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE DNA-binding protein (Fragment).
GN E2A DBP.
OS Bovine adenovirus type 2 (Mastadenovirus bos2).
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=114429;
RN [1]
RP SEQUENCE FROM N.A.
RA Harrach B., Evans P., Rusvai M., Banrevi A., Letchworth G.J.,
RA Benko M.;
RT "Codon usage differences among the protease genes of the
RT phylogenetically closely related bovine adenovirus type 1, 2 and 3.";
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U44124; AAB16762.1; -.
DR InterPro; IPR005376; Vir_DNA_Zn_bind.
DR Pfam; PF03728; Vir_DNA_Zn_bind; 1.
KW DNA-binding.
FT NON_TER 1
SQ SEQUENCE 30 AA; 3569 MW; 346B8345E6D3C41A CRC64;

Query Match 17.0%; Score 34.5; DB 12; Length 30;
Best Local Similarity 60.0%; Pred. No. 1.3e+03;
Matches 9; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

Qy 1 PKYVKQNTLKATGK 15
: | | || | | | |
Db 9 PQYRVQNTL-LPTGQ 22
: | | || | | | |

RESULT 8
Q56234
ID Q56234 PRELIMINARY; PRT; 34 AA.
AC Q56234;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Orf.
OS Thermus aquaticus.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Thermales;

```

```

OC Thermaceae; Thermus.
OX NCBI_TaxID=271;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=YT-1;
RX MEDLINE=92041736; PubMed=1939005;
RA Kirino H., Oshima T.;
RT "Molecular Cloning and nucleotide sequence of 3-Isopropylmalate
RT dehydrogenase gene (leuB) from an extreme thermophile, thermus
RT aquaticus YT-1.";
RL J. Biochem. 109:852-857(1991).
DR EMBL; D10700; BAA01543.1; -.
SQ SEQUENCE 34 AA; 3371 MW; 72449FDDC5D7B986 CRC64;

Query Match 17.0%; Score 34.5; DB 2; Length 34;
Best Local Similarity 34.8%; Pred. No. 1.5e+03;
Matches 8; Conservative 3; Mismatches 3; Indels 9; Gaps 1;

Qy 17 GPKYVKQNTLKATGKGGKGVIIIGI 39
||| ||| ||| ||| : | :
Db 21 GPK-----ETGNKGIRMGV 34
||| ||| ||| ||| : | :

RESULT 9
Q9NJN4
ID Q9NJN4 PRELIMINARY; PRT; 33 AA.
AC Q9NJN4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE DNA dependent RNA polymerase beta subunit' (Fragment).
GN RPOC1.
OS Neospora caninum.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Elmeriida; Sarcocystidae;
OC Neospora.
OX NCBI_TaxID=29176;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NC1;
RX MEDLINE=20074141; PubMed=10608442;
RA Lang-Unnasch N., Aiello D.P.;
RT "Sequence evidence for an altered genetic code in the Neospora caninum
RT plastid.";
RL Int. J. Parasitol. 29:1557-1562(1999).
DR EMBL; AF138960; AAF14262.1; -.
FT NON_TER 33
SQ SEQUENCE 33 AA; 3910 MW; 8CE8220DAF11F4EE CRC64;

Query Match 16.7%; Score 34; DB 5; Length 33;
Best Local Similarity 28.0%; Pred. No. 1.7e+03;
Matches 7; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

Qy 3 YVKQNTLKATCKGPKYVKQNTLK 27
| : | | | : | : : | |
Db 5 YIKNNTIGFRLSLASPNLIKWSLK 29
: | : | | | : | : : | |

RESULT 10
Q9KML6
ID Q9KML6 PRELIMINARY; PRT; 36 AA.
AC Q9KML6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical protein VCA0321.
GN VCA0321.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;

```

```
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.K., Mekalanos J.J., Venter J.C.,
RA Fraser C.N.;
RA "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
DR EMBL; AE004370; AAF96229.1; -.
DR TIGR; VCA0321; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 36 AA; 4073 MW; 7F9F53BF0FF9BDB4 CRC64;

Query Match 16.7%; Score 34; DB 16; Length 36;
Best Local Similarity 47.1%; Pred. No. 1.9e+03;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 9 LKLTATGKGGPKYVKONT 25
: || : || ||| |
Db 18 MILASTRTPTKAVKQAT 34

RESULT 11
Q9RI37 PRELIMINARY; PRT; 39 AA.
AC Q9RI37;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Cyclophilin A (Fragment).
GN PPIA OR CYPA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Ola;
RX MEDLINE=20422670; PubMed=10964515;
RA Colgan J., Asmal M., Luban J.;
RT "Isolation, characterization and targeted disruption of mouse Ppia:
RT cyclophilin A is not essential for mammalian cell viability.";
RL Genomics 68:167-178(2000).
DR EMBL; AF171073; AAD50996.1; -.
DR HSSP; P05092; 2CPL.
DR MGD; MGI:97749; Ppia.
DR InterPro; IPR002130; CSA_PP1ase.
DR Pfam; PF00160; pro-isomerase; 1.
DR PROSITE; PS50072; CSA_PP1ASE_2; 1.
FT NON_TER 1
FT NON_TER 39
SQ SEQUENCE 39 AA; 4324 MW; CB53F70E1092889C CRC64;

Query Match 16.5%; Score 33.5; DB 11; Length 39;
Best Local Similarity 47.4%; Pred. No. 2.4e+03;
Matches 9; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

Qy 16 KGPYVKQNTLKLATGKKG 34
: || : || ||| |
Db 4 KVPK-TAENFRALSTGEKG 21

RESULT 12
O87604 PRELIMINARY; PRT; 28 AA.
AC O87604;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Vanillate:corrinoid protein methyltransferase (Fragment).
GN ODMB.1
```

```
OS Acetobacterium dehalogenans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Eubacteriaceae; Acetobacterium.
OX NCBI_TaxID=82116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99041574; PubMed=9826201;
RA Kaufmann F., Wohlfarth G., Diekert G.;
RT "O-demethylase from Acetobacterium dehalogenans. cloning, sequencing,
RT and active expression of the gene encoding the corrinoid protein.";
RL Eur. J. Biochem. 257:515-521(1998).
DR EMBL; AF087018; AAC83696.1; -.
KW Methyltransferase; Transferase.
FT NON_TER 28
SQ SEQUENCE 28 AA; 3271 MW; EA190576604CE999 CRC64;

Query Match 16.3%; Score 33; DB 2; Length 28;
Best Local Similarity 31.6%; Pred. No. 1.9e+03;
Matches 6; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Qy 5 KONTLKLATGKGGPKYVKQ 23
: || : || ||| |
Db 5 RQNLVEVMKGGNPDREVKQ 23

RESULT 13
Q9RQ54 PRELIMINARY; PRT; 30 AA.
ID Q9RQ54
AC Q9RQ54;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ATP-dependent DNA helicase (Fragment).
GN REP.
OS Buchnera aphidicola.
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=9;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20022990; PubMed=10555290;
RA Clark M.A., Moran N.A., Baumann P.;
RT "Sequence evolution in bacterial endosymbionts having extreme base
RT compositions.";
RL Mol. Biol. Evol. 16:1586-1598(1999).
DR EMBL; AF130812; AAF13800.1; -.
DR HSSP; P09980; IUAA.
DR InterPro; IPR000212; UvrD-helicase.
DR Pfam; PF00580; UvrD-helicase; 1.
KW Helicase.
FT NON_TER 30
SQ SEQUENCE 30 AA; 3033 MW; FCOE1A84A0D95B3E CRC64;

Query Match 16.3%; Score 33; DB 2; Length 30;
Best Local Similarity 31.6%; Pred. No. 2.1e+03;
Matches 6; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Qy 20 YVKQNTLKLATGKGGVIIG 38
: : || ||| |
Db 5 FAQKNAVKLITGPCIILAG 23

RESULT 14
Q9KR17 PRELIMINARY; PRT; 31 AA.
ID Q9KR17
AC Q9KR17;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical protein VC1830.
GN VC1830.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
```

```
RN
RP
RC SEQUENCE FROM N.A.
RX STRAIN=EL TOR N16961 / SEROTYPE O1;
RA MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae";
RL Nature 406:477-483(2000).
DR EMBL; AE004259; AAF94978.1; -.
DR TIGR; VC1830; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 31 AA; 3454 MW; 62E7CBAE41EAD483 CRC64;

Query Match 16.3%; Score 33; DB 16; Length 31;
Best Local Similarity 36.4%; Pred. No. 2.1e+03;
Matches 8; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

QY 13 TGKKGPKYVKQNTLKLATGKKG 34
Db 8 THKKGKKEGRNRVYSTDGKDG 29

RESULT 15
Q9ZELL
ID Q9ZELL PRELIMINARY; PRT; 37 AA.
AC Q9ZELL;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE Hypothetical 4.3 kda protein.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pl;
RA Hofreuter D.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Pl;
RX MEDLINE=98326821; PubMed=9663688;
RA Hofreuter D., Odenbreit S., Henke G., Haas R.;
RT "Natural competence for DNA transformation in Helicobacter pylori:
RT Identification and genetic characterization of the comB locus.";
RL Mol. Microbiol. 28:1027-1038(1998).
DR EMBL; AJ123366; CAA10654.1; -.
KW Hypothetical protein.
SQ SEQUENCE 37 AA; 4325 MW; FBE3B2D972D3915B CRC64;

Query Match 16.3%; Score 33; DB 2; Length 37;
Best Local Similarity 40.9%; Pred. No. 2.6e+03;
Matches 9; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 8 TLKLTATGKKGPKYVKONTLKLKLA 29
Db 16 TSKVHEMKKSPCTLYENRLNLA 37

Search completed: December 4, 2002, 13:12:21
Job time : 30 secs
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 13:07:31 ; Search time 34 Seconds  
(without alignments)  
156.765 Million cell updates/sec

Title: US-09-362-731A-4

Perfect score: 203

Sequence: 1 PKYVKQNTLKLATGKGPYVYKQNTLKLATGKGVIGIK 40

Scoring table:

Blosum62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 398495

Minimum DB seq length: 0

Maximum DB seq length: 40

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

1: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.\*  
2: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.\*  
3: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.\*  
4: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.\*  
5: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.\*  
6: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.\*  
7: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.\*  
8: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.\*  
9: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.\*  
10: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.\*  
11: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.\*  
12: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.\*  
13: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.\*  
14: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.\*  
15: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.\*  
16: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.\*  
17: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.\*  
18: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.\*  
19: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.\*  
20: /SID22/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.\*  
21: /SID22/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.\*  
22: /SID22/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.\*  
23: /SID22/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	203	100.0	40	21 AAY82635	Influenza A virus
2	80.5	39.7	19	21 AAY99172	HLA class II bindi
3	77	37.9	19	21 AAB13854	Influenza virus ha
4	77	37.9	23	15 AAR60857	Influenza virus T-
5	77	37.9	23	15 AAR60860	Polyoxime COSM com
6	77	37.9	23	19 AAW69273	Haemagglutinin hea
7	77	37.9	24	19 AAW69268	Haemagglutinin hea
8	77	37.9	25	7 AAP60885	Synthetic peptide
9	77	37.9	25	19 AAW68315	MHC binding peptid
10	77	37.9	25	19 AAW63054	Influenza A haemag

11	77	37.9	25	20 AAY29714	Influenza virus an
12	77	37.9	25	21 AAY89679	Core polypeptide f
13	77	37.9	25	21 AAY68192	Altered MHC determ
14	77	37.9	25	21 AAY52846	Altered MHC determ
15	77	37.9	25	22 ABB01087	Viral DP178/107-11
16	77	37.9	25	22 ABB02536	Viral core polypep
17	77	37.9	25	22 AAU13633	DP178-like/DP107-1
18	77	37.9	25	22 AAB78080	Core polypeptide T
19	77	37.9	25	22 AAB58607	Altered MHC determ
20	77	37.9	26	7 AAP60887	Synthetic peptide
21	73	36.0	16	21 AAY99053	HLA class II bindi
22	73	36.0	18	23 AAE22829	Influenza virus ha
23	72	35.5	14	16 AAR79117	Peptide FHA307-320
24	72	35.5	14	19 AAW69272	Haemagglutinin hea
25	72	35.5	15	11 AAR06413	Replacement peptid
26	72	35.5	15	21 AAY99301	HLA class II bindi
27	72	35.5	16	23 AA014950	Influenza virus vi
28	72	35.5	18	21 AAY99063	HLA class II bindi
29	72	35.5	24	22 AAB15594	Peptide 13 contain
30	69	34.0	22	8 AAP70719	Equine influenza v
31	67	33.0	15	22 AAB15592	Peptide 12 for gen
32	66	32.5	13	14 AAR33500	T helper peptide i
33	66	32.5	13	14 AAR46507	Influenza Virus ha
34	66	32.5	13	15 AAR49314	HA position 307-31
35	66	32.5	13	16 AAR87434	Human MHC class II
36	66	32.5	13	16 AAR78922	Influenza 307-319
37	66	32.5	13	16 AAR70914	Influenza epitope
38	66	32.5	13	16 AAR74152	Peptide HA 307-313
39	66	32.5	13	17 AAR88362	Influenza haemaggl
40	66	32.5	13	18 AAW29188	HA 307-319 peptide
41	66	32.5	13	19 AAW54716	Peptide from Infl
42	66	32.5	13	19 AAW54674	Peptide from HA 30
43	66	32.5	13	19 AAW54729	Peptide from HA p3
44	66	32.5	13	19 AAW50104	Pan DR binding pep
45	66	32.5	13	19 AAW53600	Peptide encoded by

#### ALIGNMENTS

##### RESULT 1

AAV82635  
ID AAY82635 standard; peptide; 40 AA.

XX AAY82635;

XX AC  
XX DT 07-AUG-2000 (first entry)

XX DE Influenza A virus T cell epitope and Der p1 B cell epitope peptide.

XX KW T cell epitope; B cell epitope; allergen; antigenic;  
KW antiallergic; antiasthmatic; antiinflammatory; dermatological;  
KW immunosuppressive; vaccine; rhinitis; sinusitis; bronchial asthma;  
KW atopic dermatitis; acute urticaria; chronic urticaria;  
KW gastro-intestinal syndrome; food allergen; oro-pharyngeal syndrome;  
KW anaphylactic reaction; drug hypersensitivity; allergic reaction.

XX OS Dermatophagoides pteronyssinus.

XX OS Influenza virus.

XX OS Synthetic.

XX PN WO200006694-A2.

XX PD 10-FEB-2000.

XX PF 20-JUL-1999; 99WO-BE00092.

XX PR 30-JUL-1998; 98EP-0870167.

XX PA (UNIO ) UCB SA.

XX PI Saint-Remy J, Jacquemin M;

XX XX

DR WPI; 2000-422470/36.

PT New compound for prevention and treatment of allergies comprises at

PT least one allergen antigenic determinant recognized by a B cell and at

PT least one antigenic determinant which does not trigger T cell

PT activation -

XX

PS Claim 8; Page 35; 50pp; English.

XX

CC The present invention describes a compound (I) for the prevention and/or

CC treatment of allergy. The compound comprises at least one allergen

CC antigenic determinant (i) recognised by a B cell or an antibody secreted

CC by a B cell of a non-atopic individual and at least one antigenic

CC determinant (ii) different from the allergen that triggers T cell

CC activation. (i) has anti-allergic, antiasthmatic, antiinflammatory,

CC dermatological and immunosuppressive activities, and can be used in a

CC vaccine. (i) may be used in a pharmaceutical or cosmetic medicament to

CC treat and/or prevent allergies or a disease of allergic origin,

CC especially hypersensitivities. These include rhinitis, sinusitis,

CC bronchial asthma, atopic dermatitis, some forms of acute and chronic

CC urticaria, gastro-intestinal syndromes associated with the ingestion of

CC food allergens, oro-pharyngeal syndrome, anaphylactic reactions

CC associated with drug hypersensitivities and/or a mixture of these. The

CC use of (i) in the treatment of allergic conditions avoids the need for

CC drug treatment, which often causes undesirable side-effects. Also, prior

CC art drug therapies alleviate symptoms, but do not influence their

CC causes, however (i) actually combats the cause of an allergic reaction.

CC The present sequence represents a specifically claimed compound peptide

CC sequence from the present invention.

XX

SQ Sequence 40 AA;

Query Match 100.0%; Score 203; DB 21; Length 40;

Best Local Similarity 100.0%; Pred. No. 1.1e-20;

Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKYVKQNTLKLATGKGGPKYVKQNTLKLATGKGGVIGIK 40

DB 1 PKYVKQNTLKLATGKGGPKYVKQNTLKLATGKGGVIGIK 40

RESULT 2

AAAY99172

ID AAAY99172 standard; Peptide; 19 AA.

XX

AC AAAY99172;

XX

DT 07-AUG-2000 (first entry)

XX

DE HLA class II binding antigen epitope peptide #361.

XX

KW Human leucocyte antigen; HLA class II; antigen epitope; pharmaceutical;

KW immune response; chronic viral disease; cancer; autoimmune disease;

KW rheumatoid arthritis; multiple sclerosis; myasthenia gravis; AIDS;

KW allograft rejection; allergy; Lyme disease; hepatitis; prostate cancer;

KW glomerulonephritis; food hypersensitivity; malaria.

XX

OS Unidentified.

XX

PN WO9961916-A1.

XX

PD 02-DEC-1999.

XX

PF 28-MAY-1999; 99WO-US12066.

XX

PR 29-MAY-1998; 98US-0087192.

XX

PA (EPTM-) EPIMUNE INC.

XX

PI Sette A, Southwood S, Sidney J;

XX

DR WPI; 2000-097143/08.

XX

PT New compositions containing immunogenic peptide epitopes for various

PT HLA class II DR molecules useful for inducing helper T cell response -

PS Claim 1; Page 46; 60pp; English.

XX

CC The present invention relates to a new pharmaceutical composition

CC comprising a unit dose form of a peptide, or analogue, comprising an

CC epitope selected from those represented by peptides AAY9812-Y9939

CC which are derived from various antigens for various human leucocyte

CC antigen class DR molecules, representative of the world wide population.

CC The peptide/analogue binds to an HLA class II molecule at an IC-50 of

CC less than or equal to 1,000 nM. The pharmaceutical can be used to induce

CC a helper T cell response. The pharmaceutical focuses the immune response

CC towards selected determinants and could therefore be used in cases of

CC chronic viral diseases and cancer. Examples of diseases that can be

CC treated using the peptide containing pharmaceutical include autoimmune

CC diseases (rheumatoid arthritis, multiple sclerosis, and myasthenia

CC gravis), allograft rejection, allergies, Lyme disease, hepatitis,

CC post-streptococcal endocarditis or glomerulonephritis and food

CC hypersensitivities. The peptide epitopes can be used to enhance immune

CC responses against other immunogens administered with the peptides.

CC Diseases which can be treated using immunogenic mixtures include

CC prostate cancer, hepatitis B, hepatitis C, AIDS, renal carcinoma,

CC cervical carcinoma, lymphoma, and condyloma acuminatum. The peptides may

CC also be used to make monoclonal antibodies useful as potential diagnostic

CC or therapeutic agents. The peptides may also be useful as diagnostic

CC reagents, for example, to determine the susceptibility of an individual

CC to a treatment regimen. Also, the peptides may be used to predict which

CC individuals will be at substantial risk of developing chronic infection.

CC The selection of appropriate T and B cell epitopes should allow the

CC development of epitope based vaccines particularly towards conserved

CC epitopes of pathogens which are characterized by high sequence

CC variability such as HIV, HCV and Malaria.

XX

SQ Sequence 19 AA;

Query Match 39.7%; Score 80.5; DB 21; Length 19;

Best Local Similarity 73.1%; Pred. No. 0.00025;

Matches 19; Conservative 0; Mismatches 0; Indels 7; Gaps 1;

QY 3 YVKQNTLKLATGKGGPKYVKQNTLKL 28

DB 1 YVKQNTLKL-----YVKQNTLKL 19

RESULT 3

AAAB13854

ID AAAB13854 standard; peptide; 19 AA.

XX

AC AAAB13854;

XX

DT 13-NOV-2000 (first entry)

XX

DE Influenza virus haemagglutinin epitope 306-324.

KW Haemagglutinin; vaccine; antiviral; influenza virus; shelf-life;

KW epitope.

XX

OS Influenza virus.

XX

PN WO200032228-A2.

XX

PD 08-JUN-2000.

XX

PF 28-NOV-1999; 99WO-IL00640.

XX

PR 30-NOV-1998; 98IL-0127331.

XX

PA (YEDA ) YEDA RES & DEV CO LTD.

XX

PI Arnon R, Ben-Yedidia T, Levi R;

XX

DR WPI; 2000-412160/35.

XX Human synthetic peptide-based influenza vaccine comprises epitopes of  
PT influenza virus expressed in Salmonella flagellin -  
XX Claim 3; Page 4; 36pp; English.  
XX The present invention relates to a human synthetic peptide-based  
CC influenza vaccine. The vaccine contains at least 4 epitopes of the  
CC influenza virus each expressed individually in Salmonella flagellin.  
CC The present sequence is the T-helper epitope from influenza virus  
CC haemagglutinin epitope 306-324. The invention also includes epitopes  
CC for use in a vaccine for non-Caucasian populations. The peptide  
CC vaccine is advantageous because it has a longer shelf life, can  
CC be targeted towards specific epitopes and used to prepare a  
CC multipathogen vaccine. As the epitopes are derived from conserved  
CC regions in the virus proteins they can induce cross-strain protection.  
XX  
SQ Sequence 19 AA;  
Query Match 37.9%; Score 77; DB 21; Length 19;  
Best Local Similarity 83.3%; Pred. No. 0.00075;  
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 18 PKYVKONTLKLATGKGV 35  
DB 2 PKYVKONTLKLATGMRNV 19  
RESULT 4  
AAR60857  
ID AAR60857 standard; peptide; 23 AA.  
XX  
AC AAR60857;  
XX  
DT 05-JUN-1995 (first entry)  
XX  
DE Influenza virus T-cell epitope.  
XX  
KW Polyoxime; homopolyoxime; heteropolyoxime; peptide presentation;  
KW cell imaging; complementary orthogonal specifically active molecule;  
KW COSM; baseplate; immunogen; influenza virus; T-cell; T-lymphocyte.  
OS Synthetic.  
XX  
XX WO9425071-A.  
PN  
XX  
XX 10-NOV-1994.  
PD  
XX  
PF 05-MAY-1994; 94WO-IB00093.  
XX  
XX 05-MAY-1993; 93US-0057594.  
PR  
XX 31-AUG-1993; 93US-0105904.  
PR  
XX 31-AUG-1993; 93US-0114877.  
XX  
PA (OFFO/) OFFORD R E.  
PA (ROSE/) ROSE K.  
XX  
XX Offord RE, Rose K;  
PI  
XX  
XX WPI; 1994-357918/44.  
DR  
XX  
XX Homo- and hetero-polyoxime compounds and their preparation - used  
PT for peptide presentation to antibodies and in cell imaging etc.  
XX  
XX Disclosure; Page 60; 85pp; English.  
XX  
XX Peptides given in AAR60833-62 are used as baseplates and COSMs for  
CC the preparation of polyoximes having varying spacing, charge,  
CC lipophilicity, valency, conformational restraints, solubility and  
CC other physical and biological properties. The peptide given in  
CC AAR60857 corresponds to the T-cell epitope of influenza virus and  
CC was used to exemplify the making of polyoximes.  
XX

SQ Sequence 23 AA;  
Query Match 37.9%; Score 77; DB 15; Length 23;  
Best Local Similarity 83.3%; Pred. No. 0.00094;  
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 18 PKYVKONTLKLATGKGV 35  
DB 1 PKYVKONTLKLATGMRNV 18  
RESULT 5  
AAR60860  
ID AAR60860 standard; peptide; 23 AA.  
XX  
AC AAR60860;  
XX  
DT 05-JUN-1995 (first entry)  
XX  
DE Polyoxime COSM component.  
XX  
KW Polyoxime; homopolyoxime; heteropolyoxime; peptide presentation;  
KW cell imaging; complementary orthogonal specifically active molecule;  
KW COSM; baseplate; immunogen; influenza virus; T-cell; T-lymphocyte.  
OS Synthetic.  
XX  
XX  
FH Key Location/Qualifiers  
FT Modified-site 1  
FT /label= OTHER  
FT /note= "acetylaspartic acid"  
FT Modified-site 2  
FT /label= OTHER  
FT /note= "(carboxamidomethyl)cysteine"  
FT Modified-site 13  
FT /note= "C-terminal NHCH2CH2NH-aminooxylacetyl"  
XX  
XX WO9425071-A.  
PN  
XX  
XX 10-NOV-1994.  
PD  
XX  
PF 05-MAY-1994; 94WO-IB00093.  
XX  
XX 05-MAY-1993; 93US-0057594.  
PR  
XX 31-AUG-1993; 93US-0105904.  
PR  
XX 31-AUG-1993; 93US-0114877.  
XX  
PA (OFFO/) OFFORD R E.  
PA (ROSE/) ROSE K.  
XX  
XX Offord RE, Rose K;  
PI  
XX  
XX WPI; 1994-357918/44.  
DR  
XX  
XX Homo- and hetero-polyoxime compounds and their preparation - used  
PT for peptide presentation to antibodies and in cell imaging etc.  
XX  
XX Disclosure; Page 62; 85pp; English.  
XX  
XX Peptides given in AAR60833-62 are used as baseplates and COSMs for  
CC the preparation of polyoximes having varying spacing, charge,  
CC lipophilicity, valency, conformational restraints, solubility and  
CC other physical and biological properties. Immunogenic polyoximes  
CC were obtained using the peptides given in AAR60833-54 and AAR60858-61.  
XX  
SQ Sequence 23 AA;  
Query Match 37.9%; Score 77; DB 15; Length 23;  
Best Local Similarity 83.3%; Pred. No. 0.00094;  
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 18 PKYVKONTLKLATGKGV 35  
DB 1 PKYVKONTLKLATGMRNV 18

Db 1 PRYVKONTLKLATGMRNV 18

# RESULT 6

AAW69273  
ID AAW69273 standard; peptide; 23 AA.

XX AC AAW69273;

XX DT 29-OCT-1998 (first entry)

XX DE Haemagglutinin heavy chain (HAL) fragment.

XX KW Acryloylated peptide polymer; immune response; peptide epitope;  
KW synthetic vaccine; enzymatically cleavable site.

XX OS Influenza virus.

XX FH Key Location/Qualifiers

FT Misc-difference 1 /note= "linked to acryloylated peptide polymer"

XX FT W09834968-A1.

XX PN 13-AUG-1998.

XX PD 10-FEB-1998; 98WO-AU00076.

XX PF 03-OCT-1997; 97CA-2217321.

XX PR 11-FEB-1997; 97AU-0005071.

XX PA (CSTR ) COMMONWEALTH SCI & IND RES ORG.

XX PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.

XX PA (CSLC-) CSL LTD.

XX PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.

XX PA (UYME ) UNIV MELBOURNE.

XX PI Brandt ER, Brown LE, Ede NJ, Good MF, Jackson DC;

XX PI O'Brien-Simpson NM, Zeng W;

XX DR WPI; 1998-447177/38.

XX ACryloylated peptide polymers - useful for synthetic vaccine  
PT technology, for raising an immune response to peptide epitope and as  
PT diagnostic tool

XX PS Example 1; Page 20; 77pp; English.

XX This sequence represents a fragment of the heavy chain (HAL) of the  
CC haemagglutinin of influenza virus. This sequence was used to test the  
CC acryloylated peptide polymer of the invention. The peptide polymers are  
CC used to raise an immune response to a peptide epitope (such as this  
CC sequence), and also as diagnostic tools. Polymers (molecular wt. >600  
CC kDa.) can be prepared with virtually any number of the same or different  
CC epitopes by a method that allows purification of the individual  
CC determinants, avoids errors inherent in long sequential syntheses in  
CC which protected peptide fragments are not used, thus avoiding solubility  
CC and purification problems. Multiple copies of many different peptide  
CC epitopes may be incorporated into a single polymeric structure to allow  
CC utilisation of the range of T cell epitopes required for outbred  
CC populations in conjunction with epitopes representing different  
CC pathogenic serodemes, thus making them a significant advance in synthetic  
CC vaccine technology.

XX SQ Sequence 23 AA;

Query Match 37.9%; Score 77; DB 19; Length 23;

Best Local Similarity 83.3%; Pred. No. 0.00094;

Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 18 PRYVKONTLKLATGKKGV 35

DB 1 PRYVKONTLKLATGMRNV 18

# RESULT 7

AAW69268

XX ID AAW69268 standard; peptide; 24 AA.

XX AC AAW69268;

XX DT 29-OCT-1998 (first entry)

XX DE Haemagglutinin heavy chain (HAL) fragment.

XX KW Acryloylated peptide polymer; immune response; peptide epitope;  
KW synthetic vaccine; enzymatically cleavable site.

XX OS Influenza virus.

XX FH Key Location/Qualifiers

FT Misc-difference 1 /note= "linked to acryloylated peptide polymer"

XX FT W09834968-A1.

XX PN 13-AUG-1998.

XX PD 10-FEB-1998; 98WO-AU00076.

XX PF 03-OCT-1997; 97CA-2217321.

XX PR 11-FEB-1997; 97AU-0005071.

XX PA (CSTR ) COMMONWEALTH SCI & IND RES ORG.

XX PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.

XX PA (CSLC-) CSL LTD.

XX PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.

XX PA (UYME ) UNIV MELBOURNE.

XX PI Brandt ER, Brown LE, Ede NJ, Good MF, Jackson DC;

XX PI O'Brien-Simpson NM, Zeng W;

XX DR WPI; 1998-447177/38.

XX ACryloylated peptide polymers - useful for synthetic vaccine  
PT technology, for raising an immune response to peptide epitope and as  
PT diagnostic tool

XX PS Example 1; Page 19; 77pp; English.

XX This sequence represents a fragment of the heavy chain (HAL) of the  
CC haemagglutinin of influenza virus. This sequence was used to test the  
CC acryloylated peptide polymer of the invention. The peptide polymers are  
CC used to raise an immune response to a peptide epitope (such as this  
CC sequence), and also as diagnostic tools. Polymers (molecular wt. >600  
CC kDa.) can be prepared with virtually any number of the same or different  
CC epitopes by a method that allows purification of the individual  
CC determinants, avoids errors inherent in long sequential syntheses in  
CC which protected peptide fragments are not used, thus avoiding solubility  
CC and purification problems. Multiple copies of many different peptide  
CC epitopes may be incorporated into a single polymeric structure to allow  
CC utilisation of the range of T cell epitopes required for outbred  
CC populations in conjunction with epitopes representing different  
CC pathogenic serodemes, thus making them a significant advance in synthetic  
CC vaccine technology.

XX SQ Sequence 24 AA;

Query Match 37.9%; Score 77; DB 19; Length 24;

Best Local Similarity 83.3%; Pred. No. 0.00099;

Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 18 PRYVKONTLKLATGKKGV 35

DB 2 PRYVKONTLKLATGMRNV 19

RESULT 8  
 AAP60885  
 ID AAP60885 standard; protein; 25 AA.  
 XX  
 AC AAP60885;  
 XX  
 DT 03-OCT-2002 (updated)  
 DT 22-MAY-1991 (first entry)  
 XX  
 DE Synthetic peptide which elicits anti-influenza virus antibodies.  
 XX  
 KW Synthetic peptide; influenza virus; antibodies; diagnosis;  
 XX  
 OS Influenza virus strain X-47.  
 XX  
 PN US4625015-A.  
 XX  
 PD 25-NOV-1986.  
 XX  
 PF 29-AUG-1983; 83US-0527401.  
 XX  
 PR 29-AUG-1983; 83US-0527401.  
 PR 23-AUG-1982; 82US-0410455.  
 XX  
 PA (SCRI-) SCRIPPS CLINIC & RE.  
 XX  
 PI Green N, Alexander S;  
 XX  
 DR WPI; 1986-331839/50.  
 XX  
 PT New synthetic peptide cpds. - capable of eliciting prodn. of  
 PT antibodies that neutralise influenza virus strains  
 XX  
 PS Disclosure; page 5; 13pp; English.  
 XX  
 CC The peptide is capable of eliciting the prodn. of antibodies that  
 CC neutralise more than one strain of influenza virus. The antibodies  
 CC are useful in the treatment and prevention of influenza. The peptide  
 CC may be used in diagnosis.  
 CC (Updated on 03-OCT-2002 to add missing OS field.)  
 XX  
 SQ Sequence 25 AA;  
 Query Match 37.9%; Score 77; DB 7; Length 25;  
 Best Local Similarity 83.3%; Pred. No. 0.001;  
 Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 18 PKYVKQNTLKLATGKGV 35  
 | | | | | | | | | | | | | | | | | |  
 Db 2 PKYVKQNTLKLATGMRNV 19

RESULT 9  
 AAW68315  
 ID AAW68315 standard; peptide; 25 AA.  
 XX  
 AC AAW68315;  
 XX  
 DT 14-OCT-1998 (first entry)  
 XX  
 DE MHC binding peptide H 305-329.  
 XX  
 KW Antigen; major histocompatibility complex; MHC; lymphocyte; detection;  
 KW immobilisation; cytotoxic T-cell; tumour; leukaemia; lymphoma;  
 KW viral infection.  
 XX  
 OS Synthetic.  
 OS Influenza virus.  
 XX  
 PN W09744667-A2.  
 XX  
 PD 27-NOV-1997.

XX 21-MAY-1997; 97WO-FR00892.  
 XX  
 PR 21-MAY-1996; 96US-0651925.  
 XX  
 PA (INRM ) INSM INST NAT SANTE & RECH MEDICALE.  
 PA (INSP ) INSM PASTEUR.  
 XX  
 PI Abastado J, Kourilsky P, Langlade-Demoyen P, Lone Y;  
 XX  
 DR WPI; 1998-018653/02.  
 XX  
 PT Detection, purification and elimination of antigen-specific  
 PT lymphocytes - for producing cytotoxic T cells for immuno-therapy of  
 PT cancers and viral infection  
 XX  
 PS Disclosure; Page 27; 222pp; French.  
 XX  
 CC Peptides AAW68301-W68384 are examples of antigens (Ag) which can be  
 CC loaded onto recombinantly produced major histocompatibility complex  
 CC (MHC) molecules in a method of detecting antigen-specific lymphocytes.  
 CC The MHC-antigen complex is then immobilised on a solid support and a  
 CC sample containing cells recognising the MHC-Ag complex may be isolated.  
 CC This peptide is derived from amino acids 305-329 of the influenza virus  
 CC haemagglutinin protein. A similar method is used to isolate, purify or  
 CC eliminate Ag-specific T-cells or to produce Ag-specific cytotoxic  
 CC T-cells (CTC). The method is also used to detect and quantify  
 CC tumour-specific T-cells and to generate CTC for specific killing of  
 CC tumour cells (solid tumours, leukaemia or lymphoma) by injection into a  
 CC human or animal, but also for treating viral infections.  
 XX  
 SQ Sequence 25 AA;  
 Query Match 37.9%; Score 77; DB 19; Length 25;  
 Best Local Similarity 83.3%; Pred. No. 0.001;  
 Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 18 PKYVKQNTLKLATGKGV 35  
 | | | | | | | | | | | | | | | | | |  
 Db 2 PKYVKQNTLKLATGMRNV 19

RESULT 10  
 AAW63054  
 ID AAW63054 standard; peptide; 25 AA.  
 XX  
 AC AAW63054;  
 XX  
 DT 07-OCT-1998 (first entry)  
 XX  
 DE Influenza A haemagglutinin peptide H.305-329.  
 XX  
 KW Superantigen; treatment; cancer; tumour-specific antigen;  
 KW autoimmune disease related antigen; infection; bacterial; viral;  
 KW eukaryotic; autoimmune disease; inhibit; pathological response;  
 KW immune response.  
 XX  
 OS Synthetic.  
 OS Influenza virus.  
 XX  
 PN W09826747-A2.  
 XX  
 PD 25-JUN-1998.  
 XX  
 PF 17-DEC-1997; 97WO-US23637.  
 XX  
 PR 17-APR-1997; 97US-0044074.  
 PR 17-DEC-1996; 96US-0033172.  
 XX  
 PA (TERM/) TERMAN D S.  
 XX  
 PI Terman DS;  
 XX

DR WPI; 1998-362497/31.  
 PT Conjugates and polymers containing superantigen and therapeutic  
 PT antigen - for treatment of cancer, infection, autoimmune disease and  
 PT graft rejection, also treatment by administering lymphocytes treated  
 PT in vitro by these antigens  
 XX  
 PS Example 2; Page 40; 139pp; English.  
 CC  
 CC Synthetic peptides AAW63049-85 are used, with superantigens, to  
 CC exemplify the invention. The specification describes a method for  
 CC treatment of cancer which comprises incubating lymphocytes with  
 CC a tumour-specific antigen or autoimmune disease related antigen and  
 CC a superantigen. The treated cells are then introduced into the patient.  
 CC The superantigen and the tumour-specific antigen or autoimmune disease  
 CC related antigen can be conjugated together. The products are used  
 CC to treat cancer (carcinoma, melanoma, lymphoma etc.), infections  
 CC (bacterial, viral or eukaryotic) and autoimmune disease (e.g. idiopathic  
 CC thrombocytopenic purpura, rheumatoid arthritis, systemic lupus  
 CC erythematosus, multiple sclerosis etc.). The antigens either induce an  
 CC immune response or inhibit a pathological response.  
 XX  
 SQ Sequence 25 AA;  
 Query Match 37.9%; Score 77; DB 19; Length 25;  
 Best Local Similarity 83.3%; Pred. No. 0.001;  
 Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 18 PKYVKQNTLKLATGKKGKGV 35  
 ||||| :  
 Db 2 PKYVKQNTLKLATGMRNV 19  
 RESULT 11  
 AAY29714  
 ID AAY29714 standard; Protein; 25 AA.  
 XX  
 AC AAY29714;  
 XX  
 DT 08-NOV-1999 (first entry)  
 XX  
 DE Influenza virus antigen Infl polypeptide hapten.  
 XX  
 KW Human hepatitis B core protein; Hbc; modified; immunodominant;  
 KW nucleocapsid protein; vaccine; T cell epitope.  
 XX  
 OS Influenza virus.  
 XX  
 PN WO9940934-A1.  
 XX  
 PD 19-AUG-1999.  
 XX  
 PF 11-FEB-1999; 99WO-US03055.  
 XX  
 PR 12-FEB-1998; 98US-0074537.  
 XX  
 PA (IMMU-) IMMUNE COMPLEX CORP.  
 XX  
 PI Birkett AJ;  
 XX  
 DR WPI; 1999-527340/44.  
 XX  
 PT Conjugate of hepatitis B core protein, modified to increase  
 PT reactivity with hapten, used to raise antibodies against the hapten,  
 PT e.g. in vaccines  
 XX  
 PS Example 3; Page 38; 128pp; English.  
 CC  
 CC The present invention describes a conjugate (A) comprising a  
 CC strategically modified hepatitis B core (Hbc) protein (I) attached to  
 CC a hapten, where (I) includes amino acids (aa) 10-140 of the wild type  
 CC Hbc 193 aa sequence (given in AAY29674) and additionally has an insert  
 CC (II) in the region corresponding to aa's 50-100, where the insert is

CC of 1 to about 40 aa's and contains a chemically reactive aa residue  
 CC linked to the hapten. A vaccine containing (A), optionally in the form  
 CC of particles, is used to induce a protective antibody response against  
 CC the pathogen from which the hapten is derived, in humans or other  
 CC animals. These pathogens may be bacteria, viruses, rickettsia or  
 CC protozoa. Insertion of (II) overcomes the low reactivity of aa side  
 CC chains in native Hbc protein, increasing the reactivity with hapten  
 CC and resulting in conjugates of improved immunogenicity. Modified Hbc  
 CC can be derivatised in the form of particles by well-defined chemical  
 CC methods, and is unlikely to cause immunological side-effects. AAY29675  
 CC to AAY29735 represent polypeptide haptens used in an example from the  
 CC present invention.  
 XX  
 SQ Sequence 25 AA;  
 Query Match 37.9%; Score 77; DB 20; Length 25;  
 Best Local Similarity 83.3%; Pred. No. 0.001;  
 Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 18 PKYVKQNTLKLATGKKGKGV 35  
 ||||| :  
 Db 2 PKYVKQNTLKLATGMRNV 19  
 RESULT 12  
 AAY89679  
 ID AAY89679 standard; peptide; 25 AA.  
 XX  
 AC AAY89679;  
 XX  
 DT 23-MAY-2000 (first entry)  
 XX  
 DE Core polypeptide fragment T No. 1241.  
 XX  
 KW Retrovirus; hybrid polypeptide; enhancer; gp41; envelope protein; HIV-1;  
 KW HIV-2; SIV; pharmacokinetic; half-life; growth factor; cytokine; viral;  
 KW anti-fusogenic; differentiation factor; interleukin; interferon;  
 KW colony stimulating factor; hormone; angiogenic factor.  
 XX  
 OS Unidentified.  
 XX  
 PN WO9959615-A1.  
 XX  
 PD 25-NOV-1999.  
 XX  
 PF 20-MAY-1999; 99WO-US11219.  
 XX  
 PR 20-MAY-1998; 98US-0082279.  
 XX  
 PA (TRIM-) TRIMERIS INC.  
 XX  
 PI Barney S, Guthrie KI, Merutka G, Anwer MK, Lambert DM;  
 XX  
 DR WPI; 2000-136792/12.  
 XX  
 PT A new hybrid polypeptide with enhanced pharmacokinetic properties  
 PT comprises enhancer sequence -  
 XX  
 PS Disclosure; Page 42; 124pp; English.  
 XX  
 CC The invention relates to hybrid polypeptides comprising enhancer peptide  
 CC sequence linked to core polypeptides. The enhancer polypeptides are  
 CC derived from various retroviral envelope (gp41) protein sequences,  
 CC especially from HIV-1, HIV-2 and SIV. The enhancer peptides enhance the  
 CC pharmacokinetic properties such as increasing the half-life of any core  
 CC polypeptide that they are linked to. The core polypeptides are any  
 CC polypeptide that may be introduced into a living system and that can  
 CC function as a pharmacologically useful peptide for the treatment or  
 CC prevention of a disease. The core polypeptides are bioactive peptides  
 CC selected from a growth factor, cytokine, differentiation factor,  
 CC interleukin, interferon, colony stimulating factor, hormone or  
 CC angiogenic factor. The peptides of the invention can be used for  
 CC inhibiting viral infection and can be used in anti-viral and

CC anti-fusogenic treatments. Sequences AAY8651-Y9055 represent core  
 CC polypeptide fragments that can be used in the invention. Some sequences  
 CC among those indicated also comprise enhancer fragments at terminal ends  
 CC and form hybrid polypeptides.

XX  
 SQ Sequence 25 AA;  
 Query Match 37.9%; Score 77; DB 21; Length 25;  
 Best Local Similarity 83.3%; Pred. No. 0.001; Mismatches 1; Indels 0; Gaps 0;  
 Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 18 PKYVKONTLKLATGKGV 35  
 ||||| : I  
 Db 2 PKYVKONTLKLATGMRNV 19

RESULT 13  
 AAY68192  
 ID AAY68192 standard; Peptide: 25 AA.  
 XX  
 AC AAY68192;  
 XX  
 DT 13-APR-2000 (first entry)  
 XX  
 DE Altered MHC determinant binding peptide SEQ ID NO:24.

XX MHC class I; major histocompatibility complex; microglobulin; antigen;  
 KW immune response; immunisation; AIDS; multiple sclerosis; toxic shock;  
 KW cancer; lupus erythematosus; snake bite; cytostatic; antiviral;  
 KW immunomodulatory; dermatological; immunosuppressive; antiinflammatory;  
 KW neuroprotective.

XX Homo sapiens.  
 OS  
 XX US6011146-A.  
 PN  
 XX 04-JAN-2000.  
 PD  
 XX 07-JUN-1995; 95US-0481985.  
 PF  
 XX 15-NOV-1991; 91US-0792473.  
 PR  
 XX 05-DEC-1991; 91US-0801818.  
 PR  
 XX (INSP ) INST PASTEUR.  
 PA (INRM ) INST NAT SANTE & RECH MEDICALE.  
 PA  
 XX Kourilsky P, Mottez E, Abastado J;  
 PI  
 XX WPI; 2000-125951/11.  
 DR  
 XX

XX New recombinant DNA encoding covalently linked form of major  
 PT histocompatibility complex Class I determinant, used for immune system  
 PT stimulation, e.g. for treating cancer  
 PT  
 XX Disclosure; Column 11; 88pp; English.

XX The present invention describes a recombinant DNA molecule (I)  
 CC containing a sequence (Ia) that encodes an altered MHC (major  
 CC histocompatibility complex ) Class I determinant (II) comprises a  
 CC polypeptide with alpha1, alpha2, alpha3 and beta2-microglobulin  
 CC domains, in which alpha3 and beta2 are covalently linked, thorough C-  
 CC terminus respectively, via a nucleotide spacer sequence encoding a  
 CC polypeptide. (II) includes an antigen-binding site and when (II) and  
 CC the antigen are associated they are recognized by a mammalian T cell  
 CC receptor (TCR). (I) are used to produce (II) which are used to study  
 CC functional interactions between the various MHC domains. They can also  
 CC be used to modulate (in vivo or in vitro) the immune system by inducing  
 CC an effector response (cytotoxicity, antibody synthesis, phagocytosis)  
 CC of immune system cells, typically for treating, or immunising against;  
 CC cancer, acquired immune deficiency syndrome, lupus erythematosus,  
 CC multiple sclerosis, toxic shock and snake bite, but also for selective  
 CC destruction of autoreactive cells, diagnostically to assay T cell  
 CC receptors and to raise specific antibodies (useful for diagnosis,

CC therapy, studying MHC-associated cellular processes and for affinity  
 CC purification). AAY57558 and AAY68186 to AAY68316 are sequences used in  
 CC the exemplification of the present invention.

XX  
 SQ Sequence 25 AA;  
 Query Match 37.9%; Score 77; DB 21; Length 25;  
 Best Local Similarity 83.3%; Pred. No. 0.001; Mismatches 1; Indels 0; Gaps 0;  
 Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 18 PKYVKONTLKLATGKGV 35  
 ||||| : I  
 Db 2 PKYVKONTLKLATGMRNV 19

RESULT 14  
 AAY52846  
 ID AAY52846 standard; Peptide: 25 AA.  
 XX  
 AC AAY52846;  
 XX  
 DT 14-FEB-2000 (first entry)  
 XX  
 DE Altered MHC determinant binding peptide SEQ ID NO:24.

XX Major histocompatibility complex; MHC class I; MHC class II; antigen;  
 KW immune response; diagnosis; antibody; immunisation; autoimmune disease;  
 KW acquired immune deficiency syndrome; AIDS; cytostatic; dermatological;  
 KW anti-inflammatory; neuroprotective; immunosuppressive; antithyroid;  
 KW vaccine; lupus erythematosus; multiple sclerosis; thyroiditis;  
 KW toxic shock; tumour; snakebite.

XX Synthetic.  
 OS  
 XX Homo sapiens.  
 PN  
 XX US5976551-A.  
 XX 02-NOV-1999.  
 PD  
 XX 07-JUN-1995; 95US-0484905.  
 PF  
 XX 05-DEC-1991; 91US-0801818.  
 PR  
 XX 15-NOV-1991; 91US-0792473.  
 PR  
 XX (INSP ) INST PASTEUR.  
 PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
 PA  
 XX Kourilsky P, Mottez E, Abastado J;  
 PI  
 XX WPI; 2000-037081/03.

XX Composition containing an antigen and altered major histocompatibility  
 PT Class II determinant, used to immunize against autoimmune diseases,  
 PT e.g. acquired immune deficiency syndrome  
 PT  
 XX Claim 8; Column 11; 96pp; English.

XX The present invention describes a composition capable of eliciting  
 CC anti-major histocompatibility (MHC) antibodies. The composition  
 CC comprises an antigen associated with an altered MHC Class II determinant  
 CC (I) comprising alpha1, alpha2, beta1 and beta2 polypeptide domains  
 CC encoded by a mammalian MHC Class II locus covalently linked to form a  
 CC polypeptide (I) containing beta2, alpha2, alpha1 and beta1 domains in  
 CC sequence. The resulting Antigen-MHC complex is recognizable by the T cell  
 CC receptor. The compositions are used for immunisation against, or  
 CC treatment of, a wide range of autoimmune diseases, e.g. acquired immune  
 CC deficiency syndrome (AIDS), lupus erythematosus, multiple sclerosis,  
 CC thyroiditis, toxic shock, tumour and snakebite, depending on the nature  
 CC of antigen. (I) is also used to analyse functional interactions between  
 CC the various domains and for targeting lymphocyte receptors. Antibodies  
 CC against (I) are produced by usual methods of immunisation or cell fusion,  
 CC and may be humanised by standard methods. These antibodies are useful for  
 CC diagnosis (detection or purification of MHC gene products), therapy

CC (neutralising MHC on cell surfaces) and in the study of MHC and cellular  
 CC processes. AA233240 to AA233242 and AA252840 to AA252970 represent  
 CC sequences used in the exemplification of the present invention.

XX Sequence 25 AA;

Query Match 37.9%; Score 77; DB 21; Length 25;

Best Local Similarity 83.3%; Pred. No. 0.001;

Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 18 PKYVKQNTLKLATGKKGK 35

DB 2 PKYVKQNTLKLATGMRNV 19

RESULT 15

ABB01087

ID ABB01087 standard; Peptide; 25 AA.

XX AC ABB01087;

XX 03-JAN-2002 (first entry)

XX DE Viral DP178/107-like region peptide T1241.

XX KW Human immunodeficiency virus; HIV; respiratory syncytial virus; RSV;  
 KW virucide; heptad repeat region; transmembrane protein; gp41; HR1; HR2;  
 KW infection.

XX OS Virididae.

XX FH Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal is substituted by Ac"

FT Modified-site 25 /note= "C-terminal amide"

XX PN WO200164013-A2.

XX PD 07-SEP-2001.

XX PF 07-FEB-2001; 2001WO-US03988.

XX PR 29-FEB-2000; 2000US-0515965.

XX PA (TRIM-) TRIMERIS INC.

XX PI Antczak JB, Delmedico MK, Erickson JB, Lambert DM, Sista P;

XX DR WPI; 2001-514829/56.

XX PT Heptad repeat region peptide analogs useful for inhibiting virus/cells  
 PT fusion, useful for treating HIV and Respiratory Syncytial Virus  
 PT infection -

XX PS Disclosure; Page 55; 587pp; English.

XX CC The invention relates to isolated analogues of the heptad repeat region  
 CC peptides DP178 and DP107. DP178 and DP107 correspond to amino acids  
 CC 638-673 (heptad repeat region HR2) and 558-595 (heptad repeat region  
 CC HR1) respectively, of HIV-1/IIAI transmembrane protein gp41. The HR1  
 CC and HR2 regions of proteins interact non-covalently with each other  
 CC and/or with peptides derived from them. This interaction is required for  
 CC normal infectivity of viruses such as RSV and HIV. The heptad  
 CC repeat region peptide analogues may be used to inhibit respiratory  
 CC syncytial virus (RSV) infection in a cell. They may also be used to  
 CC inhibit HIV infection. The present sequence is a peptide provided in  
 CC the specification.

XX SQ Sequence 25 AA;

Query Match 37.9%; Score 77; DB 22; Length 25;

Best Local Similarity 83.3%; Pred. No. 0.001;

Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 18 PKYVKQNTLKLATGKKGK 35

DB 2 PKYVKQNTLKLATGMRNV 19

Search completed: December 4, 2002, 13:11:28

Job time : 35 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 13:12:27 : Search time 10 Seconds  
(without alignments)  
64.969 Million cell updates/sec

Title: US-09-362-731A-4  
Perfect score: 203  
Sequence: 1 PKYVKQNTLKLATGKPKYVKQNTLKLATGKKGVIIGIK 40

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 45477

Minimum DB seq length: 0  
Maximum DB seq length: 40

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : Published Applications\_AA:\*
- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
  - 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
  - 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
  - 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
  - 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
  - 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
  - 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	77	37.9	23	9	US-10-044-034-3
2	77	37.9	24	9	US-10-044-034-11
3	72	35.5	14	9	US-10-044-034-15
4	72	35.5	14	10	US-09-202-077-13
5	72	35.5	14	10	US-09-202-077-15
6	68	33.5	16	9	US-10-044-034-6
7	66	32.5	13	10	US-09-848-164-8
8	66	32.5	13	10	US-09-768-872-2
9	66	32.5	13	10	US-09-245-487B-23
10	66	32.5	13	10	US-09-756-983-8
11	66	32.5	13	10	US-09-756-983-14
12	63	31.0	17	9	US-10-044-034-5
13	62	30.5	13	9	US-09-987-137-7
14	61	30.0	12	9	US-10-044-034-4
15	50	24.6	10	9	US-10-044-034-27
16	42	20.7	32	10	US-09-864-761-42297
17	39	19.2	34	10	US-09-848-834A-13
18	37	18.2	29	10	US-09-732-754-1
19	35.5	17.5	29	10	US-09-570-515-19

20	35.5	17.5	31	10	US-09-864-761-39727	Sequence 39727, A
21	34	16.7	25	10	US-09-864-761-34299	Sequence 34299, A
22	34	16.7	25	10	US-09-864-761-35217	Sequence 35217, A
23	34	16.7	26	10	US-09-864-761-34286	Sequence 34286, A
24	34	16.7	29	10	US-09-864-761-48516	Sequence 48516, A
25	34	16.7	39	10	US-09-864-761-38971	Sequence 38971, A
26	33.5	16.5	25	10	US-09-810-310-43	Sequence 43, Appl
27	33	16.3	7	9	US-10-044-034-13	Sequence 13, Appl
28	33	16.3	20	8	US-08-981-824-4	Sequence 4, Appl
29	33	16.3	21	10	US-09-853-830-54	Sequence 54, Appl
30	32.5	16.0	26	10	US-09-864-761-41444	Sequence 41444, A
31	32	15.8	21	10	US-09-853-830-57	Sequence 57, Appl
32	32	15.8	26	10	US-09-864-761-45578	Sequence 45578, A
33	32	15.8	26	10	US-09-973-451-21	Sequence 21, Appl
34	32	15.8	31	10	US-09-848-834A-15	Sequence 15, Appl
35	32	15.8	31	10	US-09-983-019-6	Sequence 6, Appl
36	32	15.8	40	10	US-09-879-957-82	Sequence 82, Appl
37	31.5	15.5	22	10	US-09-269-533A-5	Sequence 5, Appl
38	31.5	15.5	28	10	US-09-864-761-45448	Sequence 45448, A
39	31.5	15.5	38	10	US-09-766-050-2	Sequence 2, Appl
40	31	15.3	21	10	US-09-853-830-56	Sequence 56, Appl
41	31	15.3	26	10	US-09-764-877-1384	Sequence 1384, Ap
42	31	15.3	28	10	US-09-989-903-50	Sequence 50, Appl
43	31	15.3	33	10	US-09-864-761-44501	Sequence 44501, A
44	31	15.3	35	10	US-09-864-866-47	Sequence 47, Appl
45	30.5	15.0	25	10	US-09-864-761-38703	Sequence 38703, A

ALIGNMENTS

RESULT 1  
US-10-044-034-3  
; Sequence 3, Application US/10044034  
; Patent No. US20020169264A1  
; GENERAL INFORMATION:  
; APPLICANT: JACKSON, DAVID C.  
; APPLICANT: O'BRIEN-SIMPSON, NEIL M.  
; APPLICANT: BROWN, LORENA E.  
; APPLICANT: EDE, NICHOLAS J.  
; APPLICANT: BRANDT, EVELYN R.  
; APPLICANT: GOOD, MICHAEL F.  
; TITLE OF INVENTION: POLYMERS INCORPORATING PEPTIDES  
; FILE REFERENCE: FBRC:006  
; CURRENT APPLICATION NUMBER: US/10/044,034  
; CURRENT FILING DATE: 2002-01-11  
; PRIOR APPLICATION NUMBER: P05071  
; PRIOR FILING DATE: 1997-02-11  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 23  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Peptides  
US-10-044-034-3

Query Match 37.9%; Score 77; DB 9; Length 23;  
Best Local Similarity 83.3%; Pred. No. 9.9e-05;  
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 18 PKYVKQNTLKLATGKKGKV 35  
Db 1 PKYVKQNTLKLATGMRNV 18  
|||||  
|||||

RESULT 2  
US-10-044-034-11  
; Sequence 11, Application US/10044034  
; Patent No. US20020169264A1  
; GENERAL INFORMATION:

; APPLICANT: JACKSON, DAVID C.  
; APPLICANT: O'BRIEN-SIMPSON, NEIL M.  
; APPLICANT: BROWN, LORENA E.  
; APPLICANT: EDE, NICHOLAS J.  
; APPLICANT: BRANDT, EVELYN R.  
; APPLICANT: GOOD, MICHAEL F.  
; TITLE OF INVENTION: POLYMERS INCORPORATING PEPTIDES  
; FILE REFERENCE: FBRC:006  
; CURRENT APPLICATION NUMBER: US/10/044,034  
; CURRENT FILING DATE: 2002-01-11  
; PRIOR APPLICATION NUMBER: P05071  
; PRIOR FILING DATE: 1997-02-11  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 24  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-10-044-034-11

Query Match 37.9%; Score 77; DB 9; Length 24;  
Best Local Similarity 83.3%; Pred. No. 0.0001;  
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 18 PKYVKONTLKLATGKKGV 35  
| | | | | | | | | | | | | | | | | |  
Db 2 PKYVKONTLKLATGMRNV 19

RESULT 3  
US-10-044-034-15  
; Sequence 15, Application US/10044034  
; Patent No. US20020169264A1  
; GENERAL INFORMATION:  
; APPLICANT: JACKSON, DAVID C.  
; APPLICANT: O'BRIEN-SIMPSON, NEIL M.  
; APPLICANT: BROWN, LORENA E.  
; APPLICANT: EDE, NICHOLAS J.  
; APPLICANT: BRANDT, EVELYN R.  
; APPLICANT: GOOD, MICHAEL F.  
; TITLE OF INVENTION: POLYMERS INCORPORATING PEPTIDES  
; FILE REFERENCE: FBRC:006  
; CURRENT APPLICATION NUMBER: US/10/044,034  
; CURRENT FILING DATE: 2002-01-11  
; PRIOR APPLICATION NUMBER: P05071  
; PRIOR FILING DATE: 1997-02-11  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-10-044-034-15

Query Match 35.5%; Score 72; DB 9; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00026;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKYVKONTLKLATG 14  
| | | | | | | | | | | | | | | |  
Db 1 PKYVKONTLKLATG 14

RESULT 4  
US-09-202-077-13  
; Sequence 13, Application US/09202077  
; Patent No. US20020103335A1

; GENERAL INFORMATION:  
; APPLICANT: OLDHAM, Keith  
; APPLICANT: EDWARDS, Philip N.  
; APPLICANT: LUKE, Richard W.  
; APPLICANT: COTTON, Ronald  
; TITLE OF INVENTION: PEPTIDE DERIVATIVES  
; FILE REFERENCE: 1991-140  
; CURRENT APPLICATION NUMBER: US/09/202,077  
; CURRENT FILING DATE: 1999-05-26  
; PRIOR APPLICATION NUMBER: GB 9611881.5  
; PRIOR FILING DATE: 1996-06-07  
; PRIOR APPLICATION NUMBER: GB 9622890.3  
; PRIOR FILING DATE: 1996-11-02  
; PRIOR APPLICATION NUMBER: GB 9701491  
; PRIOR FILING DATE: 1997-06-03  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: NP\_BIND  
; LOCATION: (1)  
; OTHER INFORMATION: Biotin-Ahx-  
; NAME/KEY: NP\_BIND  
; LOCATION: (14)  
; OTHER INFORMATION: -OH  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: peptide  
US-09-202-077-13

Query Match 35.5%; Score 72; DB 10; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00026;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKYVKONTLKLATG 14  
| | | | | | | | | | | | | | | |  
Db 1 PKYVKONTLKLATG 14

RESULT 5  
US-09-202-077-15  
; Sequence 15, Application US/09202077  
; Patent No. US20020103335A1  
; GENERAL INFORMATION:  
; APPLICANT: OLDHAM, Keith  
; APPLICANT: EDWARDS, Philip N.  
; APPLICANT: LUKE, Richard W.  
; APPLICANT: COTTON, Ronald  
; TITLE OF INVENTION: PEPTIDE DERIVATIVES  
; FILE REFERENCE: 1991-140  
; CURRENT APPLICATION NUMBER: US/09/202,077  
; CURRENT FILING DATE: 1999-05-26  
; PRIOR APPLICATION NUMBER: GB 9611881.5  
; PRIOR FILING DATE: 1996-06-07  
; PRIOR APPLICATION NUMBER: GB 9622890.3  
; PRIOR FILING DATE: 1996-11-02  
; PRIOR APPLICATION NUMBER: GB 9701491  
; PRIOR FILING DATE: 1997-06-03  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-202-077-15

Query Match 35.5%; Score 72; DB 10; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00026;



Db 1 PKYVKONTLKLAT 13  
|||||  
RESULT 9  
US-09-245-487B-23  
; Sequence 23, Application US/09245487B  
; Patent No. US2002003878/Al  
; GENERAL INFORMATION:  
; APPLICANT: STROMINGER, Jack L.  
; FALK, Kirsten  
; ROTZSCHKE, Olaf  
; TITLE OF INVENTION: MHC Binding Peptide Oligomers and  
; Methods of Use  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Testa, Hurwitz & Thibault  
; STREET: 125 High St.  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/245,487B  
; FILING DATE: 05-Feb-1999  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: TWOMEY, Michael J.  
; REGISTRATION NUMBER: 38,349  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 23:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 13 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:  
US-09-245-487B-23

Query Match 32.5%; Score 66; DB 10; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.0016;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PKYVKONTLKLAT 13  
Db 1 PKYVKONTLKLAT 13  
|||||

RESULT 10  
US-09-756-983-8  
; Sequence 8, Application US/09756983  
; Patent No. US20020122818A1  
; GENERAL INFORMATION:  
; APPLICANT: Albani, Salvatore  
; TITLE OF INVENTION: METHODS FOR ISOLATION, QUANTIFICATION,  
; CHARACTERIZATION AND MODULATION OF  
; TITLE OF INVENTION: ANTIGEN-SPECIFIC T CELLS  
; FILE REFERENCE: 246/285-CIP  
; CURRENT APPLICATION NUMBER: US/09/756,983  
; CURRENT FILING DATE: 2001-01-09  
; PRIOR APPLICATION NUMBER: 60/105,018  
; PRIOR FILING DATE: 1998-10-20  
; PRIOR APPLICATION NUMBER: 09/421,506  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR APPLICATION NUMBER: PCT/US99/2466  
; PRIOR FILING DATE: 1999-10-19

; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthesized peptide derived from the influenza virus  
US-09-756-983-8

Query Match 32.5%; Score 66; DB 10; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.0016;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PKYVKONTLKLAT 13  
Db 1 PKYVKONTLKLAT 13  
|||||

RESULT 11  
US-09-756-983-14  
; Sequence 14, Application US/09756983  
; Patent No. US20020122818A1  
; GENERAL INFORMATION:  
; APPLICANT: Albani, Salvatore  
; TITLE OF INVENTION: METHODS FOR ISOLATION, QUANTIFICATION,  
; CHARACTERIZATION AND MODULATION OF  
; TITLE OF INVENTION: ANTIGEN-SPECIFIC T CELLS  
; FILE REFERENCE: 246/285-CIP  
; CURRENT APPLICATION NUMBER: US/09/756,983  
; CURRENT FILING DATE: 2001-01-09  
; PRIOR APPLICATION NUMBER: 60/105,018  
; PRIOR FILING DATE: 1998-10-20  
; PRIOR APPLICATION NUMBER: 09/421,506  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR APPLICATION NUMBER: PCT/US99/2466  
; PRIOR FILING DATE: 1999-10-19  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-756-983-14

Query Match 32.5%; Score 66; DB 10; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.0016;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PKYVKONTLKLAT 13  
Db 1 PKYVKONTLKLAT 13  
|||||

RESULT 12  
US-10-044-034-5  
; Sequence 5, Application US/10044034  
; Patent No. US20020169264A1  
; GENERAL INFORMATION:  
; APPLICANT: JACKSON, DAVID C.  
; APPLICANT: O'BRIEN-SIMPSON, NEIL M.  
; APPLICANT: BROWN, LORENA E.  
; APPLICANT: EDE, NICHOLAS J.  
; APPLICANT: BRANDT, EVELYN R.  
; APPLICANT: GOOD, MICHAEL F.  
; TITLE OF INVENTION: POLYMERS INCORPORATING PEPTIDES  
; FILE REFERENCE: FBRC:006  
; CURRENT APPLICATION NUMBER: US/10/044,034  
; CURRENT FILING DATE: 2002-01-11  
; PRIOR APPLICATION NUMBER: P05071  
; PRIOR FILING DATE: 1997-02-11  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1

```
; SEQ ID NO 5
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-044-034-5

Query Match      31.0%; Score 63; DB 9; Length 17;
Best Local Similarity 92.9%; Pred. No. 0.0056;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 16 KGPVKVQNTLKL 29
DB 4 KYPVKVQNTLKL 17

RESULT 13
US-09-987-137-7
; Sequence 7, Application US/09987137
; Patent No. US20020177534A1
; GENERAL INFORMATION:
; APPLICANT: Verheijden, Gijbertus Franciscus Maria
; APPLICANT: Boots, Anna Maria Helena
; TITLE OF INVENTION: No. US20020177534A1el Peptides for use in Treatment of T-cell Med
; TITLE OF INVENTION: Cartilage Destruction in Auto-immune Diseases
; FILE REFERENCE: Verheijdenseq
; CURRENT APPLICATION NUMBER: US/09/987,137
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: 08/981,340
; PRIOR FILING DATE: 1997-12-18
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-987-137-7

Query Match      30.5%; Score 62; DB 9; Length 13;
Best Local Similarity 92.3%; Pred. No. 0.0056;
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKYVKQNTLKLAT 13
DB 1 PKYVKQNTLKLAT 13

RESULT 14
US-10-044-034-4
; Sequence 4, Application US/10044034
; Patent No. US20020169264A1
; GENERAL INFORMATION:
; APPLICANT: JACKSON, DAVID C.
; APPLICANT: O'BRIEN-SIMPSON, NEIL M.
; APPLICANT: BROWN, LORENA E.
; APPLICANT: EDE, NICHOLAS J.
; APPLICANT: BRANDT, EVELYN R.
; APPLICANT: GOOD, MICHAEL F.
; TITLE OF INVENTION: POLYMERS INCORPORATING PEPTIDES
; FILE REFERENCE: FBRC:006
; CURRENT APPLICATION NUMBER: US/10/044,034
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: P05071
; PRIOR FILING DATE: 1997-02-11
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-044-034-4

Query Match      30.0%; Score 61; DB 9; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0007;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKYVKQNTLKL 12
DB 1 PKYVKQNTLKL 12

RESULT 15
US-10-044-034-27
; Sequence 27, Application US/10044034
; Patent No. US20020169264A1
; GENERAL INFORMATION:
; APPLICANT: JACKSON, DAVID C.
; APPLICANT: O'BRIEN-SIMPSON, NEIL M.
; APPLICANT: BROWN, LORENA E.
; APPLICANT: EDE, NICHOLAS J.
; APPLICANT: BRANDT, EVELYN R.
; APPLICANT: GOOD, MICHAEL F.
; TITLE OF INVENTION: POLYMERS INCORPORATING PEPTIDES
; FILE REFERENCE: FBRC:006
; CURRENT APPLICATION NUMBER: US/10/044,034
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: P05071
; PRIOR FILING DATE: 1997-02-11
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-044-034-27

Query Match      24.8%; Score 50; DB 9; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KYVKQNTLKL 11
DB 1 KYVKQNTLKL 10

Search completed: December 4, 2002, 13:16:09
Job time : 10 secs
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 13:10:51 ; Search time 14 Seconds  
(without alignments)  
84.065 Million cell updates/sec

Title: US-09-362-731A-4

Perfect score: 203

Sequence: 1 PKYVKQNTLKLATGKPKYVKQNTLKLATGKGVIIIGIK 40

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 174272

Minimum DB seq length: 0

Maximum DB seq length: 40

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued\_Patents\_AA:\*
- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
  - 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
  - 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
  - 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*
  - 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	77	37.9	23	4	US-08-537-928-25
2	77	37.9	25	2	US-08-484-905-24
3	77	37.9	25	3	US-08-481-985B-24
4	77	37.9	25	4	US-08-370-476-24
5	77	37.9	25	4	US-09-248-588-85
6	77	37.9	25	4	US-09-082-279B-1063
7	77	37.9	25	4	US-09-315-304B-1063
8	72	35.5	14	3	US-09-125-517A-61
9	72	35.5	14	4	US-09-319-870A-9
10	72	35.5	14	4	US-09-308-175A-21
11	72	35.5	15	3	US-09-125-517A-60
12	72	35.5	16	4	US-09-284-625-24
13	72	35.5	16	4	US-09-284-625-26
14	71	35.0	25	4	US-08-952-877-7
15	70	34.5	23	4	US-08-537-928-26
16	66	32.5	13	1	US-07-831-895C-5
17	66	32.5	13	1	US-08-305-871A-1
18	66	32.5	13	1	US-08-465-167A-22
19	66	32.5	13	2	US-08-480-190-24
20	66	32.5	13	2	US-08-596-387B-8
21	66	32.5	13	2	US-08-488-379-24
22	66	32.5	13	2	US-08-968-676-164
23	66	32.5	13	4	US-09-055-475-1
24	66	32.5	13	4	US-09-067-615-8
25	66	32.5	13	4	US-08-464-496-18
26	66	32.5	13	4	US-09-194-285-53
27	66	32.5	13	4	US-08-788-822A-2

28	66	32.5	13	4	US-08-197-484-99	Sequence 99, Appl
29	66	32.5	13	4	US-08-627-820-22	Sequence 22, Appl
30	66	32.5	13	5	PCT-US92-07218-18	Sequence 18, Appl
31	66	32.5	13	5	PCT-US93-07545-20	Sequence 24, Appl
32	66	32.5	13	5	PCT-US94-10257A-20	Sequence 20, Appl
33	66	32.5	13	5	PCT-US95-02121-99	Sequence 99, Appl
34	66	32.5	13	5	PCT-US95-04121-53	Sequence 53, Appl
35	66	32.5	13	5	PCT-US95-09816A-8	Sequence 8, Appl
36	66	32.5	14	4	US-09-009-953-2	Sequence 2, Appl
37	65	32.0	14	4	US-09-319-870A-8	Sequence 8, Appl
38	62	30.5	13	1	US-08-619-645-9	Sequence 9, Appl
39	62	30.5	13	2	US-08-634-493-9	Sequence 9, Appl
40	62	30.5	14	4	US-09-082-279B-509	Sequence 509, App
41	62	30.5	14	4	US-09-315-304B-509	Sequence 509, App
42	60	29.6	13	5	PCT-US95-04121-57	Sequence 57, Appl
43	60	29.6	13	5	PCT-US95-04121-58	Sequence 58, Appl
44	60	29.6	13	5	PCT-US95-04121-59	Sequence 59, Appl
45	59	29.1	14	4	US-09-308-175A-19	Sequence 19, Appl

#### ALIGNMENTS

##### RESULT 1

US-08-537-928-25  
; Sequence 25, Application US/08537928  
; Patent No. 6217873  
; GENERAL INFORMATION:  
; APPLICANT: Keith, Rose  
; APPLICANT: Offord, Robin E  
; TITLE OF INVENTION: POLYXIME COMPOSITIONS AND THEIR PREPARATION  
; FILE REFERENCE: GREN-001/0305  
; CURRENT APPLICATION NUMBER: US/08/537,928  
; CURRENT FILING DATE: 1996-01-05  
; EARLIER APPLICATION NUMBER: PCT/IB94/00093  
; EARLIER FILING DATE: 1994-05-05  
; EARLIER APPLICATION NUMBER: 08/105,904  
; EARLIER FILING DATE: 1993-08-31  
; EARLIER APPLICATION NUMBER: 08/114,877  
; EARLIER FILING DATE: 1993-08-31  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 25  
; LENGTH: 23  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-08-537-928-25

Query Match 37.9%; Score 77; DB 4; Length 23;  
Best Local Similarity 83.3%; Pred. No. 0.00024;  
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 18 PKYVKQNTLKLATGKGV 35  
|||||  
Db 1 PKYVKQNTLKLATGMRNV 18

##### RESULT 2

US-08-484-905-24  
; Sequence 24, Application US/08484905  
; Patent No. 5976551  
; GENERAL INFORMATION:  
; APPLICANT: Mottez, Estelle  
; APPLICANT: Abastado, Jean-Pierre  
; APPLICANT: Kourilsky, Philippe  
; TITLE OF INVENTION: An Altered Major Histocompatibility  
; TITLE OF INVENTION: Complex(MHC) Determinant and Methods for Using the  
; NUMBER OF SEQUENCES: 127  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &



```
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 05243.0001-01000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-370-476-24

Query Match 37.9%; Score 77; DB 4; Length 25;
Best Local Similarity 83.3%; Pred. No. 0.00026;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 18 PKYVKQNTLKLATGKKGV 35
  |||||
DB 2 PKYVKQNTLKLATGMRNV 19

RESULT 5
US-09-248-588-85
; Sequence 85, Application US/09248588
; Patent No. 6231864
; GENERAL INFORMATION:
; APPLICANT: Birkett, Ashley J.
; TITLE OF INVENTION: Strategically Modified Hepatitis B Core Proteins and
; FILE OF INVENTION: their Derivatives
; FILE REFERENCE: SYN-101 4564/69529
; CURRENT APPLICATION NUMBER: US/09/248,588
; CURRENT FILING DATE: 1999-02-11
; EARLIER APPLICATION NUMBER: 60/074537
; EARLIER FILING DATE: 1998-02-12
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 85
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Influenza virus
; OTHER INFORMATION: Core polypeptide
US-09-248-588-85

Query Match 37.9%; Score 77; DB 4; Length 25;
Best Local Similarity 83.3%; Pred. No. 0.00026;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 18 PKYVKQNTLKLATGKKGV 35
  |||||
DB 2 PKYVKQNTLKLATGMRNV 19

RESULT 6
US-09-082-279B-1063
; Sequence 1063, Application US/09082279B
; Patent No. 6258782
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn
; APPLICANT: Guthrie, Kelly
; APPLICANT: Merutka, Gene
; APPLICANT: Anwer, Mohamed
; APPLICANT: Lambert, Dennis
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
; FILE OF INVENTION: PHARMACOKINETIC PROPERTIES
; FILE REFERENCE: 7872-043
; CURRENT APPLICATION NUMBER: US/09/082,279B
; CURRENT FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1515
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1063
; LENGTH: 25

; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 05243.0001-01000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-370-476-24

Query Match 37.9%; Score 77; DB 4; Length 25;
Best Local Similarity 83.3%; Pred. No. 0.00026;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 18 PKYVKQNTLKLATGKKGV 35
  |||||
DB 2 PKYVKQNTLKLATGMRNV 19

RESULT 7
US-09-315-304B-1063
; Sequence 1063, Application US/09315304B
; Patent No. 6348568
; GENERAL INFORMATION:
; APPLICANT: Barney, S.
; APPLICANT: Guthrie, K.
; APPLICANT: Merutka, G.
; APPLICANT: Anwer, M.
; APPLICANT: Lambert, D.
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
; FILE OF INVENTION: PROPERTIES
; FILE REFERENCE: 7872-052
; CURRENT APPLICATION NUMBER: US/09/315,304B
; PRIOR APPLICATION NUMBER: 09/082,279
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1667
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1063
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-315-304B-1063

Query Match 37.9%; Score 77; DB 4; Length 25;
Best Local Similarity 83.3%; Pred. No. 0.00026;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 18 PKYVKQNTLKLATGKKGV 35
  |||||
DB 2 PKYVKQNTLKLATGMRNV 19

RESULT 8
US-09-125-517A-61
; Sequence 61, Application US/09125517A
; Patent No. 6087336
; GENERAL INFORMATION:
; APPLICANT: Cotton, Ronald
; APPLICANT: Edwards, Philip Neil
; APPLICANT: Luke, Richard William Arthur
; TITLE OF INVENTION: Peptide Derivatives
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.
; STREET: 555 Thirteenth St., N.W., Suite 701 East
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```



;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/125,517A  
;; FILING DATE: 20-AUG-1998  
;; CLASSIFICATION: 530  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: GB 9603855.9  
;; FILING DATE: 23-FEB-1996  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: GB 9620819.4  
;; FILING DATE: 05-OCT-1996  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Ernst, Barbara G.  
;; REGISTRATION NUMBER: 30,377  
;; REFERENCE/DOCKET NUMBER: 1991-127  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 202-783-6040  
;; TELEFAX: 202-783-6031  
;; INFORMATION FOR SEQ ID NO: 61:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 14 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-09-125-517A-61

Query Match 35.5%; Score 72; DB 3; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00065;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKYVKQNTLKLATG 14  
| | | | | | | | | | | | | | | |  
Db 1 PKYVKQNTLKLATG 14

RESULT 9  
US-09-319-870A-9  
; Sequence 9, Application US/09319870A  
; Patent No. 6184207  
; GENERAL INFORMATION:  
; APPLICANT: Astrazeneca  
; APPLICANT: Luke, Richard W. A.  
; TITLE OF INVENTION: Inhibitors of Peptide Binding to MHC Class II Proteins  
; FILE REFERENCE: 1991-175  
; CURRENT APPLICATION NUMBER: US/09/319,870A  
; CURRENT FILING DATE: 1999-06-14  
; PRIOR APPLICATION NUMBER: PCT/GB97/03397  
; PRIOR FILING DATE: 1996-12-12  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Modified  
; OTHER INFORMATION: Sequence  
US-09-319-870A-9

Query Match 35.5%; Score 72; DB 4; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00065;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKYVKQNTLKLATG 14  
| | | | | | | | | | | | | | | |  
Db 1 PKYVKQNTLKLATG 14

RESULT 10  
US-09-308-175A-21  
; Sequence 21, Application US/09308175A

;; Patent No. 6355617  
;; GENERAL INFORMATION:  
;; APPLICANT: LUKE, Richard William Arthur  
;; APPLICANT: COTTON, Ronald  
;; TITLE OF INVENTION: PEPTIDE DERIVATIVES  
;; FILE REFERENCE: 1991-174  
;; CURRENT APPLICATION NUMBER: US/09/308,175A  
;; CURRENT FILING DATE: 1999-05-17  
;; PRIOR APPLICATION NUMBER: PCT/GB97/03199  
;; PRIOR FILING DATE: 1997-11-21  
;; PRIOR APPLICATION NUMBER: GB 9624562.6  
;; PRIOR FILING DATE: 1996-11-27  
;; NUMBER OF SEQ ID NOS: 22  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 21  
;; LENGTH: 14  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-308-175A-21

Query Match 35.5%; Score 72; DB 4; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00065;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKYVKQNTLKLATG 14  
| | | | | | | | | | | | | | | |  
Db 1 PKYVKQNTLKLATG 14

RESULT 11  
US-09-125-517A-60  
; Sequence 60, Application US/09125517A  
; Patent No. 6087336  
; GENERAL INFORMATION:  
; APPLICANT: Cotton, Ronald  
; APPLICANT: Edwards, Philip Neil  
; APPLICANT: Luke, Richard William Arthur  
; TITLE OF INVENTION: Peptide Derivatives  
; NUMBER OF SEQUENCES: 62  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.  
; STREET: 555 Thirteenth St., N.W., Suite 701 East  
; CITY: Tower  
; STATE: DC  
; COUNTRY: U.S.A.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/125,517A  
; FILING DATE: 20-AUG-1998  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9603855.9  
; FILING DATE: 23-FEB-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9620819.4  
; FILING DATE: 05-OCT-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ernst, Barbara G.  
; REGISTRATION NUMBER: 30,377  
; REFERENCE/DOCKET NUMBER: 1991-127  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-783-6040  
; TELEFAX: 202-783-6031  
; INFORMATION FOR SEQ ID NO: 60:

SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1  
OTHER INFORMATION: /product= "OTHER"  
OTHER INFORMATION: /note= "biotinyl-2-aminohexanoyl"  
US-09-125-517A-60

Query Match 35.5%; Score 72; DB 3; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0007;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKYVKQNTLKLATG 14  
Db 2 PKYVKQNTLKLATG 15

## RESULT 12

US-09-284-625-24  
Sequence 24, Application US/09284625  
Patent No. 6207644

## GENERAL INFORMATION:

APPLICANT: Luke, Richard William  
TITLE OF INVENTION: Peptide Analogues Containing a 7-membered Lactam Ring  
FILE REFERENCE: 1991-169  
CURRENT APPLICATION NUMBER: US/09/284,625  
CURRENT FILING DATE: 1999-04-16  
PRIOR APPLICATION NUMBER: PCT/GB97/02837  
PRIOR FILING DATE: 1997-10-14  
PRIOR APPLICATION NUMBER: GB 96218367  
PRIOR FILING DATE: 1996-10-19  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 24  
LENGTH: 16  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: MOD\_RES  
LOCATION: (1)  
OTHER INFORMATION: Biotin-Ahx  
NAME/KEY: MOD\_RES  
LOCATION: (16)  
OTHER INFORMATION: -OH  
US-09-284-625-24

Query Match 35.5%; Score 72; DB 4; Length 16;  
Best Local Similarity 100.0%; Pred. No. 0.00076;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKYVKQNTLKLATG 14  
Db 2 PKYVKQNTLKLATG 15

## RESULT 13

US-09-284-625-26  
Sequence 26, Application US/09284625  
Patent No. 6207644

## GENERAL INFORMATION:

APPLICANT: Luke, Richard William  
TITLE OF INVENTION: Peptide Analogues Containing a 7-membered Lactam Ring  
FILE REFERENCE: 1991-169  
CURRENT APPLICATION NUMBER: US/09/284,625  
CURRENT FILING DATE: 1999-04-16  
PRIOR APPLICATION NUMBER: PCT/GB97/02837

PRIOR FILING DATE: 1997-10-14  
PRIOR APPLICATION NUMBER: GB 96218367  
PRIOR FILING DATE: 1996-10-19  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 26  
LENGTH: 16  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: MOD\_RES  
LOCATION: (1)  
OTHER INFORMATION: -H  
NAME/KEY: MOD\_RES  
LOCATION: (16)  
OTHER INFORMATION: -OH  
US-09-284-625-26

Query Match 35.5%; Score 72; DB 4; Length 16;  
Best Local Similarity 100.0%; Pred. No. 0.00076;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKYVKQNTLKLATG 14  
Db 2 PKYVKQNTLKLATG 15

## RESULT 14

US-08-992-877-7  
Sequence 7, Application US/08992877  
Patent No. 6340461

## GENERAL INFORMATION:

APPLICANT: Terman, David S  
TITLE OF INVENTION: SUPERANTIGEN BASED METHODS AND COMPOSITIONS FOR  
TITLE OF INVENTION: TREATMENT OF INFECTIOUS DISEASE  
FILE REFERENCE: superantigen  
CURRENT APPLICATION NUMBER: US/08/992,877  
CURRENT FILING DATE: 1997-12-17  
PRIOR APPLICATION NUMBER: 60/044,074  
PRIOR FILING DATE: 1997-04-17  
NUMBER OF SEQ ID NOS: 78  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 7  
LENGTH: 25  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: antigen  
US-08-992-877-7

Query Match 35.0%; Score 71; DB 4; Length 25;  
Best Local Similarity 77.8%; Pred. No. 0.0017;  
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 18 PKYVKQNTLKLATGKKG 35  
Db 2 PKYVKQNTLKLATGKGM 19

## RESULT 15

US-08-537-928-26  
Sequence 26, Application US/08537928  
Patent No. 6217873

## GENERAL INFORMATION:

APPLICANT: Keith, Rose  
APPLICANT: Offord, Robin E  
TITLE OF INVENTION: POLYOXIME COMPOSITIONS AND THEIR PREPARATION  
FILE REFERENCE: GFEN-001/03US  
CURRENT APPLICATION NUMBER: US/08/537,928  
CURRENT FILING DATE: 1996-01-05  
EARLIER APPLICATION NUMBER: PCT/IB94/00093  
EARLIER FILING DATE: 1994-05-05

; EARLIER APPLICATION NUMBER: 08/105,904  
; EARLIER FILING DATE: 1993-08-31  
; EARLIER APPLICATION NUMBER: 08/114,877  
; EARLIER FILING DATE: 1993-08-31  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 26  
; LENGTH: 23  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (1)  
; OTHER INFORMATION: Aminoxyacetyl-Pro  
US-08-537-928-26

Query Match 34.5%; Score 70; DB 4; Length 23;  
Best Local Similarity 82.4%; Pred. No. 0.0022;  
Matches 14; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 19 KYVKQNTLKLATGKGV 35

Db 2 KYVKQNTLKLATGMRNV 18

Search completed: December 4, 2002, 13:13:06  
Job time : 14 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 13:15:57 ; Search time 14 Seconds  
(without alignments)  
219.736 Million cell updates/sec

Title: US-09-362-731A-5  
Perfect score: 181  
Sequence: 1 QYIKANSKFIGITELGGCHGSEPCNIHRGKPF 32

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 6878

Minimum DB seq length: 0  
Maximum DB seq length: 32

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	37	20.4	26	1	SMNC
2	34	18.8	15	2	metallothionein -
3	34	18.8	15	2	dichloromethane de
4	34	18.8	30	2	dichloromethane de
5	34	18.8	30	2	replication initia
6	33.5	18.5	31	2	protein R10E8.7 li
7	33.5	18.5	31	2	gene x protein - h
8	33.5	18.5	31	2	gene x protein - h
9	33	18.2	23	2	pregnancy-specific
10	32.5	18.0	24	2	lethal peptide I -
11	32.5	18.0	31	2	gene x protein - h
12	32.5	18.0	31	2	gene x protein - h
13	32.5	18.0	31	2	gene x protein - h
14	32.5	18.0	31	2	gene x protein - h
15	32.5	18.0	31	2	gene x protein - h
16	32.5	18.0	31	2	gene x protein - h
17	32.5	18.0	31	2	gene x protein - h
18	32.5	18.0	31	2	gene x protein - h
19	32.5	18.0	31	2	gene x protein - h
20	32.5	18.0	31	2	gene x protein - h
21	32.5	18.0	31	2	gene x protein - h
22	32.5	18.0	31	2	gene x protein - h
23	32.5	18.0	31	2	gene x protein - h
24	32.5	18.0	31	2	gene x protein - h
25	32.5	18.0	31	2	gene x protein - h
26	32.5	18.0	31	2	gene x protein - h
27	32.5	18.0	31	2	gene x protein - h
28	32.5	18.0	31	2	gene x protein - h
29	32.5	18.0	31	2	gene x protein - h

RESULT 1  
SMNC  
metallothionein - Neurospora crassa  
C:Species: Neurospora crassa  
C:Date: 31-May-1980 #sequence\_revision 31-Mar-1993 #text\_change 22-Jun-1999  
C:Accession: A24641; A03287  
R:Muenger, K.; Germann, U.A.; Lerch, K.  
EMBO J. 4, 2665-2668, 1985  
A:Title: Isolation and structural organization of the Neurospora crassa copper metallo  
A:Reference number: A24641; MUID:86030247; PMID:2932331  
A:Accession: A24641  
A:Molecule type: DNA  
A:Residues: 1-26 <MUE>  
A:Cross-references: GB:X03009; NID:g2986; PIDN:CAA26793.1; PID:g2987  
R:Lerch, K.  
Nature 284, 368-370, 1980  
A:Title: Copper metallothionein, a copper-binding protein from Neurospora crassa.  
A:Reference number: A03287; MUID:80143244; PMID:6444697  
A:Accession: A03287  
A:Molecule type: protein  
A:Residues: 2-26 <LER>  
C:Genetics:  
A:Introns: 18/1  
C:Superfamily: metallothionein  
C:Keywords: chelation; metal binding; metal-thiolate cluster  
F:4,6,12,14,18,20,23/Binding site: transition metal ions (Cys) #status experimental  
Query Match 20.4%; Score 37; DB 1; Length 26;  
Best Local Similarity 46.2%; Pred. No. 1.6e+02;  
Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;  
QY 17 GCHGSEPCNIHRG 29  
||| : |||  
DB 5 GCSGASSCNCSSG 17

RESULT 2  
A60929  
dichloromethane dehalogenase (EC 4.5.1.3) - Hyphomicrobium sp. (ATCC 43129) (fragment  
C:Species: Hyphomicrobium sp.  
C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 07-Oct-1994  
C:Accession: A60929  
R:Kohler-Staub, D.; Hartmann, S.; Gaelli, R.; Suter, F.; Leisinger, T.  
J. Gen. Microbiol. 132, 2837-2843, 1986  
A:Title: Evidence for identical dichloromethane dehalogenases in different methylotro  
A:Reference number: A60929  
A:Accession: A60929  
A:Molecule type: protein  
A:Residues: 1-15 <ROH>  
C:Keywords: carbon-halide lyase  
Query Match 18.8%; Score 34; DB 2; Length 15;

#### ALIGNMENTS

30 32 17.7 24 2 T42846  
31 32 17.7 25 1 SMNR  
32 32 17.7 30 2 B56591  
33 32 17.7 32 2 PC4433  
34 32 17.7 32 2 B82365  
35 31.5 17.4 31 2 S53176  
36 31.5 17.4 31 2 S53190  
37 31.5 17.4 31 2 S53192  
38 31.5 17.4 31 2 S53233  
39 31.5 17.4 31 2 S53284  
40 31 17.1 26 2 S55029  
41 30.5 16.9 25 2 G85581  
42 30 16.6 24 2 I73583  
43 30 16.6 30 2 B61322  
44 30 16.6 31 2 S53153  
45 29 16.0 13 2 S74130

rePA translation p  
metallothionein -  
E75 steroid recept  
paired box transcr  
hypothetical prote  
gene x protein - h  
gene x protein - h  
gene x protein - h  
gene x protein - h  
gene x protein - h  
CAP3 protein - ant  
unknown protein en  
alpha 1-proteinase  
4K protein, hypoch  
gene x protein - h  
NADH oxidase - Gia

Best Local Similarity 75.0%; Pred. No. 2.6e+02; Mismatches 0; Gaps 0; Indels 2;

QY 23 PCNIHRGK 30  
| | | | |  
Db 5 PTNIHTGK 12

## RESULT 3

B60929  
dichloromethane dehalogenase (EC 4.5.1.3) - Pseudomonas sp. (fragment)  
C:Species: Pseudomonas sp.  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 07-Oct-1994  
C:Accession: B60929  
R:Kohler-Staub, D.; Hartmans, S.; Gaelli, R.; Suter, F.; Leisinger, T.  
J. Gen. Microbiol. 132, 2837-2843, 1986  
A:Title: Evidence for identical dichloromethane dehalogenases in different methylotrophic  
A:Reference number: A60929  
A:Contents: DSM 1565  
A:Accession: B60929  
A:Molecule type: protein  
A:Residues: 1-15 <KOH>  
C:Keywords: carbon-halide lyase

Query Match 18.8%; Score 34; DB 2; Length 15;  
Best Local Similarity 75.0%; Pred. No. 2.6e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 23 PCNIHRGK 30  
| | | | |  
Db 5 PTNIHTGK 12

## RESULT 4

S13753  
replication initiation gene A protein [similarity] - phage phi-K (fragment)  
N:Contains: DNA ligase (EC 6.5.1.-); DNA lyase (EC 4.2.99.-)  
C:Species: phage phi-K  
C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 05-May-2000  
R:Kodaira, K.I.; Miyata, T.; Suzuki, K.; Nakano, K.; Taketo, A.  
Biochim. Biophys. Acta 1008, 123-124, 1989  
A:Title: Possible finger structure in gene A protein of Microviridae.  
A:Reference number: S04334; MUID:89247441; PMID:2785817  
A:Accession: S13753  
A:Molecule type: DNA  
A:Residues: 1-30 <KOD>  
A:Cross-references: EMBL:X12610; NID:gi5096; PTDN:CAA31129.1; PID:g930130  
A:Note: the authors translated the codon CAG for residue 15 as Glu  
C:Genetics:  
A:Gene: A  
C:Superfamily: phage phi-X174 gene A protein  
C:Keywords: carbon-oxygen lyase; DNA binding; DNA replication initiation; ligase; zinc  
F:3-25/Region: zinc finger

Query Match 18.8%; Score 34; DB 2; Length 30;  
Best Local Similarity 38.5%; Pred. No. 4.9e+02;  
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 8 KFIGITELGGCHG 20  
: : : | | | |  
Db 5 RVLCPVEFGQHG 17

## RESULT 5

F89406  
protein R10E8.7 [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
C:Accession: F89406  
R:anonymus, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology

A:Reference number: A75000; MUID:99069613; PMID:9851916  
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.  
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;  
A:Accession: F89406  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-30 <STO>  
A:Cross-references: GB:chr\_V; PIDN:CA04647.1; PID:g3879068; GSPDB:GN00023; CESP:R10E  
A:Note: cDNA EST EMBL:D36612 comes from this gene  
C:Genetics:  
A:Gene: R10E8.7  
A:Map position: 5

Query Match 18.8%; Score 34; DB 2; Length 30;  
Best Local Similarity 35.0%; Pred. No. 4.9e+02;  
Matches 7; Conservative 3; Mismatches 6; Indels 4; Gaps 1;

QY 16 GGCHG-----SEPCNIHRGKP 31  
| | | | | : | : | : |  
Db 3 GCCHSLDKFSQTCELSRVRP 22

## RESULT 6

S53208  
gene X protein - hepatitis B virus (isolate patient Castag'83) (fragment)  
C:Species: hepatitis B virus, HBV  
A:Variety: isolate patient Castag'83  
C:Date: 08-Jul-1995 #sequence\_revision 03-Aug-1995 #text\_change 26-Aug-1999  
C:Accession: S53208  
R:Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.  
submitted to the EMBL Data Library, March 1995  
A:Reference number: S53112  
A:Accession: S53208  
A:Molecule type: DNA  
A:Residues: 1-31 <LAT>  
A:Cross-references: EMBL:X85290; NID:g736114; PIDN:CAA59608.1; PID:g736115  
C:Genetics:  
A:Gene: X  
C:Superfamily: hepatitis B virus gene X protein

Query Match 18.5%; Score 33.5; DB 2; Length 31;  
Best Local Similarity 38.1%; Pred. No. 5.9e+02;  
Matches 8; Conservative 1; Mismatches 7; Indels 5; Gaps 1;

QY 10 IGITELGGCH-----GSEPCN 25  
| : | | | | |  
Db 6 IKVIILGGCRHKLVCAAPCN 26

## RESULT 7

S53254  
gene X protein - hepatitis B virus (isolate patient Mannoni-1'87) (fragment)  
C:Species: hepatitis B virus, HBV  
A:Variety: isolate patient Mannoni-1'87  
C:Date: 08-Jul-1995 #sequence\_revision 03-Aug-1995 #text\_change 26-Aug-1999  
C:Accession: S53254  
R:Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.  
submitted to the EMBL Data Library, March 1995  
A:Reference number: S53112  
A:Accession: S53254  
A:Molecule type: DNA  
A:Residues: 1-31 <LAT>  
A:Cross-references: EMBL:X85309; NID:g736181; PIDN:CAA59649.1; PID:g736182  
A:Experimental source: isolate patient Mannoni-1'87  
C:Genetics:  
A:Gene: X  
C:Superfamily: hepatitis B virus gene X protein

Query Match 18.5%; Score 33.5; DB 2; Length 31;  
Best Local Similarity 43.8%; Pred. No. 5.9e+02;  
Matches 7; Conservative 0; Mismatches 4; Indels 5; Gaps 1;

QY 15 LGGCH-----GSEPCN 25

Db 11 LGGCRHKLVCAPSPCN 26  
|||||

## RESULT 8

gene X protein - hepatitis B virus (isolate patient Pische'90) (fragment)  
S53287  
C:Species: hepatitis B virus, HBV  
A:Variety: isolate patient Pische'90  
C:Date: 08-Jul-1995 #sequence\_revision 03-Aug-1995 #text\_change 26-Aug-1999  
C:Accession: S53287  
R:Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.  
submitted to the EMBL Data Library, March 1995  
A:Reference number: S53112  
A:Accession: S53287  
A:Molecule type: DNA  
A:Residues: 1-31 <LAI>  
A:Cross-references: EMBL:X85262; NID:g736228; PIDN:CAA59534.1; PID:g736229  
A:Experimental source: isolate patient Pische'90  
C:Genetics:  
A:Gene: X  
C:Superfamily: hepatitis B virus gene X protein

Query Match 18.5%; Score 33.5; DB 2; Length 31;  
Best Local Similarity 43.8%; Pred. No. 5.9e+02;  
Matches 7; Conservative 0; Mismatches 4; Indels 5; Gaps 1;

Qy 15 LGGCH-----GSEPCN 25  
|||||

Db 11 LGGCRHKLVCAPSPCN 26

## RESULT 9

A44524  
pregnancy-specific glycoprotein SBU-3-57 - sheep (fragment)  
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 31-Dec-1993  
C:Accession: A44524  
R:Atkinson, Y.H.  
submitted to the Protein Sequence Database, June 1993  
A:Reference number: A44524

A:Accession: A44524  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-23 <ATK>  
C:Keywords: glycoprotein

Query Match 18.2%; Score 33; DB 2; Length 23;  
Best Local Similarity 62.5%; Pred. No. 5.3e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 25 NIHRGKPF 32

Db 15 NIHXGNPY 22  
||| | |

## RESULT 10

B44008

lethal peptide I - Wagler's pit viper

N:Alternate names: waglerin SL-I

C:Species: Trimeresurus wagleri (Wagler's pit viper)

C:Date: 27-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 01-Mar-1996

C:Accession: B44008; A39559

R:Schmidt, J.J.; Weinstein, S.A.; Smith, L.A.

Toxinon 30, 1027-1036, 1992

A:Title: Molecular properties and structure-function relationships of lethal peptides from  
A:Reference number: A44008; MUID:93069257; PMID:1440639

A:Accession: B44008

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-24 <SCH>

A:Experimental source: venom

A:Note: sequence extracted from NCBI backbone (NCBIP:119453)

R:Weinstein, S.A.; Schmidt, J.J.; Bernheimer, A.W.; Smith, L.A.  
Toxinon 29, 227-236, 1991

A:Title: Characterization and amino acid sequences of two lethal peptides isolated from  
A:Reference number: A39559; MUID:91263101; PMID:2048140

A:Accession: A39559

A>Status: preliminary

A:Molecule type: protein

A:Residues: 3-24 <WEI>

Query Match 18.0%; Score 32.5; DB 2; Length 24;  
Best Local Similarity 45.5%; Pred. No. 6.5e+02;  
Matches 10; Conservative 2; Mismatches 7; Indels 3; Gaps 2;

Qy 11 GITELGCGCHGSEPCN-IHRGKP 31

Db 4 GKPDLRPC--PPCHYIPRPKP 23  
| : | | | : | | | |

## RESULT 11

S53132

gene X protein - hepatitis B virus (isolate patient Tarcisloc'92) (fragment)  
C:Species: hepatitis B virus, HBV

A:Variety: isolate patient Tarcisloc'92

C:Date: 08-Jul-1995 #sequence\_revision 03-Aug-1995 #text\_change 26-Aug-1999

C:Accession: S53132

R:Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.

submitted to the EMBL Data Library, March 1995

A:Reference number: S53112

A:Accession: S53132

A:Molecule type: DNA

A:Residues: 1-31 <LAI>

A:Cross-references: EMBL:X85255; NID:g736009; PIDN:CAA59515.1; PID:g736010

A:Experimental source: isolate patient Tarcisloc'92

C:Genetics:

A:Gene: X

C:Superfamily: hepatitis B virus gene X protein

Query Match 18.0%; Score 32.5; DB 2; Length 31;  
Best Local Similarity 43.8%; Pred. No. 8.2e+02;  
Matches 7; Conservative 0; Mismatches 4; Indels 5; Gaps 1;

Qy 15 LGGCH-----GSEPCN 25

Db 11 LGGCRHKLVCAPAPCN 26  
|||||

## RESULT 12

S53135

gene X protein - hepatitis B virus (isolate patient Buesca'92) (fragment)  
C:Species: hepatitis B virus, HBV

A:Variety: isolate patient Buesca'92

C:Date: 08-Jul-1995 #sequence\_revision 03-Aug-1995 #text\_change 26-Aug-1999

C:Accession: S53135

R:Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.

submitted to the EMBL Data Library, March 1995

A:Reference number: S53112

A:Accession: S53135

A:Molecule type: DNA

A:Residues: 1-31 <LAI>

A:Cross-references: EMBL:X85264; NID:g736013; PIDN:CAA59539.1; PID:g736014

A:Experimental source: isolate patient Buesca'92

C:Genetics:

A:Gene: X

C:Superfamily: hepatitis B virus gene X protein

Query Match 18.0%; Score 32.5; DB 2; Length 31;  
Best Local Similarity 43.8%; Pred. No. 8.2e+02;  
Matches 7; Conservative 0; Mismatches 4; Indels 5; Gaps 1;

Qy 15 LGGCH-----GSEPCN 25

Db 11 LGGCRHKLVCAPAPCN 26  
|||||

## RESULT 13

S53138  
gene X protein - hepatitis B virus (isolate patient Frailis'92 and others) (fragment)  
C:Species: hepatitis B virus, HBV  
A:Variety: isolate patient Frailis'92; isolate patient Dettori'85  
C:Date: 08-Jul-1995 #sequence\_revision 03-Aug-1995 #text\_change 26-Aug-1999  
C:Accession: S53138; S53226  
R:Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.  
submitted to the EMBL Data Library, March 1995  
A:Reference number: S53112  
A:Accession: S53138  
A:Molecule type: DNA  
A:Residues: 1-31 <LAI>  
A:Cross-references: EMBL:X85265; NID:g736017; PIDN:CAA59542.1; PID:g736018  
A:Experimental source: isolate patient Frailis'92  
A:Accession: S53226  
A:Molecule type: DNA  
A:Residues: 1-31 <LAW>  
A:Cross-references: EMBL:X85297; NID:g736140; PIDN:CAA59623.1; PID:g736141  
A:Experimental source: isolate patient Dettori'85  
C:Genetics:  
A:Gene: X  
C:Superfamily: hepatitis B virus gene X protein

Query Match 18.0%; Score 32.5; DB 2; Length 31;  
Best Local Similarity 43.8%; Pred. No. 8.2e+02;  
Matches 7; Conservative 0; Mismatches 4; Indels 5; Gaps 1;

QY 15 LGGCH-----GSEPCN 25  
|||||  
Db 11 LGGRHKLVCAPACN 26

## RESULT 14

S53141  
gene X protein - hepatitis B virus (isolate patient Macis'89 and others) (fragment)  
C:Species: hepatitis B virus, HBV  
A:Variety: isolate patient Macis'89; isolate patient Urtis'89; isolate patient Porcu'91;  
C:Date: 08-Jul-1995 #sequence\_revision 03-Aug-1995 #text\_change 26-Aug-1999  
C:Accession: S53141; S53170; S53185; S53230; S53248  
R:Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.  
submitted to the EMBL Data Library, March 1995  
A:Reference number: S53112  
A:Accession: S53141  
A:Molecule type: DNA  
A:Residues: 1-31 <LAI>  
A:Cross-references: EMBL:X85266; NID:g736021; PIDN:CAA59545.1; PID:g736022; EMBL:X85276;  
85299; NID:g736146; PID:g736147; EMBL:X85306; NID:g736172; PID:g736173  
A:Experimental source: isolate patient Macis'89; isolate patient Urtis'89; isolate patient  
C:Genetics:  
A:Gene: X  
C:Superfamily: hepatitis B virus gene X protein

Query Match 18.0%; Score 32.5; DB 2; Length 31;  
Best Local Similarity 43.8%; Pred. No. 8.2e+02;  
Matches 7; Conservative 0; Mismatches 4; Indels 5; Gaps 1;

QY 15 LGGCH-----GSEPCN 25  
|||||  
Db 11 LGGRHKLVCAPACN 26

## RESULT 15

S53147  
gene X protein - hepatitis B virus (isolate patient Sini'90) (fragment)  
C:Species: hepatitis B virus, HBV  
A:Variety: isolate patient Sini'90  
C:Date: 08-Jul-1995 #sequence\_revision 03-Aug-1995 #text\_change 26-Aug-1999  
C:Accession: S53147  
R:Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.  
submitted to the EMBL Data Library, March 1995  
A:Reference number: S53112

A:Accession: S53147  
A:Molecule type: DNA  
A:Residues: 1-31 <LAI>  
A:Cross-references: EMBL:X85268; NID:g736029; PIDN:CAA59551.1; PID:g736030  
A:Experimental source: isolate patient Sini'90  
C:Genetics:  
A:Gene: X  
C:Superfamily: hepatitis B virus gene X protein

Query Match 18.0%; Score 32.5; DB 2; Length 31;  
Best Local Similarity 43.8%; Pred. No. 8.2e+02;  
Matches 7; Conservative 0; Mismatches 4; Indels 5; Gaps 1;

QY 15 LGGCH-----GSEPCN 25  
|||||  
Db 11 LGGRHKLVCAPACN 26

Search completed: December 4, 2002, 13:18:05  
Job time : 14 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 13:13:12 ; Search time 11 Seconds

(without alignments)  
120.658 Million cell updates/sec

Title: US-09-362-731A-5

Perfect score: 181

Sequence: 1 QYIKANSKFIGITELGGCHGSEPCNIHRGKPF 32

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 2095

Minimum DB seq length: 0

Maximum DB seq length: 32

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	20.4	25	1 MT_NEUCR	P02807 neurospora
2	36	19.9	15	1 CX1B_CONBE	P58624 conus betul
3	36	19.9	26	1 SPRC_MUSVI	P36379 mustela vis
4	33	18.2	27	1 CX07_CONCN	P58916 conus conso
5	32.5	18.0	22	1 LPL_TRIWA	P24335 trimeresuru
6	32	17.7	25	1 MT_AGABI	P04358 agaricus bi
7	31.5	17.4	30	1 CXVA_CONER	P58782 conus ermin
8	31	17.1	26	1 MTL_COGLG	Q99334 colletotric
9	30.5	16.9	30	1 CXVB_CONER	P58783 conus ermin
10	29	16.0	21	1 NDK_CANAL	Q9ur66 candida alb
11	29	16.0	31	1 GUN2_SCLSC	P21834 sclerotinia
12	28	15.5	12	1 CXST_CONTE	P58846 conus texti
13	28	15.5	20	1 VSP1_AKGBI	P33588 agkistrodon
14	28	15.5	20	1 VSP1_TRIOK	P20005 trimeresuru
15	28	15.5	21	1 NRIA_ACISP	P33036 acinetobact
16	28	15.5	25	1 GUN4_VIBFI	P08991 vibrio fisc
17	28	15.5	28	1 GUN_SCHCO	P81190 schizophyll
18	28	15.5	28	1 TX02_AGEAP	P15971 agelenopsis
19	27.5	15.2	28	1 SCX1_LEIQH	P80669 leirus qui
20	27	14.9	25	1 CXA4_CONPU	P55963 conus purpu
21	27	14.9	30	1 CV07_VIOOD	P58439 viola odora
22	27	14.9	31	1 LC70_LACPA	P80959 lactobacill
23	26.5	14.6	22	1 LP2_TRIWA	P58930 trimeresuru
24	26	14.4	16	1 CXA1_CONEP	P56638 conus episc
25	26	14.4	22	1 PSP2_PHYPA	P80661 physcomitre
26	26	14.4	25	1 CX0B_CONMA	P05485 conus magus
27	26	14.4	27	1 TXA3_ANESI	P01535 anemonia su
28	26	14.4	28	1 SCX2_BUTSI	P15230 buthus sind
29	26	14.4	30	1 SCX2_TITSE	P08816 titus serr
30	26	14.4	31	1 PETL_OENHO	Q9mtk4 oenothera h
31	25.5	14.1	25	1 AMP3_MELGA	P80393 meleagris g
32	25	13.8	15	1 CXA2_CONAL	P56640 conus aulic
33	25	13.8	16	1 CXA1_CONAL	P56639 conus aulic

34 25 13.8 16 1 CXA3\_CONAL P56641 conus aulic  
35 25 13.8 16 1 CXAA\_CONPE P50984 conus penna  
36 25 13.8 16 1 CXAB\_CONPE P50985 conus penna  
37 25 13.8 21 1 MDH\_RHOPA P80458 rhodopseudo  
38 25 13.8 22 1 CR33\_LITCE P56240 litoria cae  
39 25 13.8 22 1 CR34\_LITCE P56241 litoria cae  
40 25 13.8 23 1 TYPH\_LACCA P19663 lactobacill  
41 25 13.8 26 1 CX07\_CONTE P56714 conus texti  
42 25 13.8 30 1 UC35\_MAIZE P80641 zea mays (m  
43 25 13.8 32 1 PHNS\_DESMU P13062 desulfovibr  
44 24.5 13.5 27 1 CH12\_STROI P29116 streptomyce  
45 24 13.3 11 1 TKC2\_CALVO P41518 calliphora

## ALIGNMENTS

RESULT 1

MT\_NEUCR

ID MT\_NEUCR STANDARD; PRT; 25 AA.

AC P02807; 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Metallothionein (MT).

GN CMT.

OS Neurospora crassa.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;

OC Sordariales; Sordariaceae; Neurospora.

OX NCBI\_TaxID=5141;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=86030247; PubMed=2932331;

RA Muenger K., Germann U.A., Lerch K.;

RT "Isolation and structural organization of the Neurospora crassa

copper metallothionein gene.";

RL EMBO J. 4:2665-2668(1985).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=91112714; PubMed=2148862;

RA Romeyer F.M., Jacobs F.A., Brousseau R.;

RT "Expression of a Neurospora crassa metallothionein and its variants

in Escherichia coli.";

RL Appl. Environ. Microbiol. 56:2748-2754(1990).

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=88029908; PubMed=2959528;

RA Muenger K., Germann U.A., Lerch K.;

RT "Isolation and regulation of expression of the Neurospora crassa

copper metallothionein gene.";

RL Experientia Suppl. 52:393-400(1987).

RN [4]

RP SEQUENCE

RX MEDLINE=80143244; PubMed=6444697;

RA Lerch K.;

RT "Copper metallothionein, a copper-binding protein from Neurospora

crassa.";

RL Nature 284:368-370(1980).

RN [5]

RP STRUCTURE BY NMR.

RX MEDLINE=89302887; PubMed=2525920;

RA Malikayil J.A., Lerch K., Armitage I.M.;

RT "Proton NMR studies of a metallothionein from Neurospora crassa:

sequence-specific assignments by NOE measurements in the rotating

frame.";

RL Biochemistry 28:2991-2995(1989).

CC -I- INDUCTION: BY COPPER.

CC -I- MISCELLANEOUS: THE SEVEN CYSTEINES BIND SIX COPPER (CUPROUS) IONS.

CC -I- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY: FAMILY 8.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its



use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
 DR EMBL; X03009; CAA26793.1; -  
 DR EMBL; M59836; AAA33594.1; -  
 DR EMBL; M27709; AAA33595.1; -  
 DR PIR; A24641; SHNC.  
 KW Metal-binding; Metal-thiolate cluster; Copper.  
 FT INIT\_MET 0 0  
 FT METAL 3 3 COPPER.  
 FT METAL 5 5 COPPER.  
 FT METAL 11 11 COPPER.  
 FT METAL 13 13 COPPER.  
 FT METAL 17 17 COPPER.  
 FT METAL 19 19 COPPER.  
 FT METAL 22 22 COPPER.  
 FT CONFLICT 21 21 N -> T (IN REF. 2).  
 SQ SEQUENCE 25 AA; 2234 MW; 256CB635FD422FC CRC64;

Query Match 20.4%; Score 37; DB 1; Length 25;  
 Best Local Similarity 46.2%; Pred. No. 46;  
 Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 17 GCHGSEPCNIHRG 29  
 || | : || |  
 DB 4 GCGASSCNCGSG 16

## RESULT 2

CX1B\_CONBE STANDARD; PRT; 15 AA.  
 AC P58624;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Conotoxin BtxIb.  
 OS Conus betulinus (Beech cone).  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
 OC Neogastropoda; Conoidea; Conidae; Conus.  
 OX NCBI\_TaxID=89764;  
 RN [1]  
 RP SEQUENCE AND MASS SPECTROMETRY.  
 RC TISSUE=Venom;  
 RX MEDLINE=20058566; PubMed=10591037;  
 RA Chen J.-S., Fan C.-X., Hu K.-P., Wei K.-H., Zhong M.-N.;  
 RT "Studies on conotoxins of Conus betulinus.";  
 RL J. Nat. Toxins 8:341-349(1999).  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.  
 CC -!- MASS SPECTROMETRY: MW=1642.5; METHOD=MALDI.  
 CC -!- SIMILARITY: BELONGS TO THE M-SUPERFAMILY OF CONOTOXINS.  
 KW Neurotoxin; Toxin.  
 FT DISULFID 1 9 PROBABLE.  
 FT DISULFID 2 12 PROBABLE.  
 FT DISULFID 6 13 PROBABLE.  
 SQ SEQUENCE 15 AA; 1650 MW; 3749B4F08E311337 CRC64;

Query Match 19.9%; Score 36; DB 1; Length 15;  
 Best Local Similarity 71.4%; Pred. No. 40;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 18 CHGSEPC 24  
 ||| ||  
 DB 6 CHGCVPC 12

## RESULT 3

SPRC\_MUSVI STANDARD; PRT; 26 AA.  
 ID SPRC\_MUSVI  
 AC P36379;  
 DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE SPARC (Secreted protein acidic and rich in cysteine) (Osteonectin)  
 DE (ON) (Basement membrane protein BM-40) (Fragment).  
 GN SPARC.  
 OS Mustela vison (American mink).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Mustellidae; Mustellinae;  
 OC Mustela.  
 OX NCBI\_TaxID=9667;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RA Ralph D., McClelland M., Welsh J.;  
 RL Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: APPEARS TO REGULATE CELL GROWTH THROUGH INTERACTIONS WITH THE EXTRACELLULAR MATRIX AND CYTOKINES. BINDS CALCIUM AND COPPER, SEVERAL TYPES OF COLLAGEN, ALBUMIN, THROMBOSPONDIN, PDGF AND CELL MEMBRANES. THERE ARE TWO CALCIUM BINDING SITES; A ACIDIC DOMAIN THAT BINDS 5 TO 8 CA++ WITH A LOW AFFINITY AND A EF-HAND LOOP THAT BINDS A CA++ ION WITH A HIGH AFFINITY.  
 CC -!- DEVELOPMENTAL STAGE: EXPRESSED AT HIGH LEVELS IN TISSUES UNDERGOING MORPHOGENESIS, REMODELING AND WOUND REPAIR.  
 CC -!- SIMILARITY: CONTAINS 1 OSTONECTIN-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
 CC EMBL; U00596; AAA16139.1; -  
 DR HSSP; P09486; LSRA.  
 DR InterPro; IPR002048; EF-hand.  
 DR InterPro; IPR001999; Osteonectin.  
 DR PROSITE; PS00018; EF HAND: PARTIAL.  
 DR PROSITE; PS00612; OSTONECTIN\_1; PARTIAL.  
 DR PROSITE; PS00613; OSTONECTIN\_2; PARTIAL.  
 KW Extracellular matrix; Basement membrane; Glycoprotein;  
 KW Calcium-binding; Copper.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 26 AA; 2940 MW; B4B45A9ACABA124A CRC64;

Query Match 19.9%; Score 36; DB 1; Length 26;  
 Best Local Similarity 33.3%; Pred. No. 67;  
 Matches 7; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 8 KFTGITELGCGHSGPCNIHR 28  
 | : | : | : | : | : | :  
 DB 2 KYALGEWAGCGTKEKDIDK 22

RESULT 4  
 CX07\_CONCN STANDARD; PRT; 27 AA.  
 AC P58916;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Omega-conotoxin CnVIIa.  
 OS Conus consors (Singed cone).  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;  
 OC Neogastropoda; Conoidea; Conidae; Conus.  
 OX NCBI\_TaxID=101297;  
 RN [1]  
 RP SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.  
 RC TISSUE=Venom;  
 RX PubMed=11724570;  
 RA Favreau P., Gilles N., Lamthanh H., Bournaud R., Shimahara T.,  
 RA Bouet F., Laboute P., Letourneux Y., Menez A., Molgo J., Le Gall F.;  
 RT "A new omega-conotoxin that targets N-type voltage-sensitive calcium

```
RT channels with unusual specificity."
RL Biochemistry 40:14567-14575(2001).
CC -!- FUNCTION: Omega-conotoxins act at presynaptic membranes, they bind
CC and block voltage-sensitive calcium channels (VSCC). This toxin
CC blocks N-type calcium channels, but unexpectedly, does not show
CC any blocking activity at amphibian neuromuscular junction. Causes
CC shaking activity, and, at higher doses, causes mild tremors when
CC injected intracerebroventricularly into mice. Causes paralysis,
CC and, at higher doses, causes death when injected intramuscularly
CC into fish.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -!- MASS SPECTROMETRY: MW=2847.74; METHOD=Electrospray.
CC -!- SIMILARITY: BELONGS TO THE O-SUPERFAMILY OF CONOTOXINS. OMEGA-TYPE
CC FAMILY.
KW Presynaptic neurotoxin; Neurotoxin; Toxin; Calcium channel inhibitor;
KW Hydroxylation; Amidation.
FT BINDING 13 13
FT FT
FT DISULFID 1 16
FT FT BY SIMILARITY.
FT DISULFID 8 20
FT FT BY SIMILARITY.
FT DISULFID 15 27
FT FT BY SIMILARITY.
FT MOD_RES 7 7
FT FT HYDROXYLATION.
FT MOD_RES 27 27
FT FT AMIDATION.
SQ SEQUENCE 27 AA; 2839 MW; B9DEFD1491F2CB4A CRC64;

Query Match 18.2%; Score 33; DB 1; Length 27;
Best Local Similarity 46.2%; Pred. No. 1.9e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 18 CHGSEPCNIHRGK 30
||||| : : : :
DB 16 CHGS--CSSSKGR 26

RESULT 5
LPL_TRIWA STANDARD; PRT; 22 AA.
AC P24335;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lethal peptide I.
OS Trimeresurus wagleri (Wagler's pit viper) (Tropidolaemus wagleri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Viperidae; Crotalinae; Tropidolaemus.
OX NCBI_TaxID=8770;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=91263101; PubMed=2048140;
RA Weinstein S.A., Schmidt J.J., Bernheimer A.W., Smith L.A.;
RT "Characterization and amino acid sequences of two lethal peptides
RT isolated from venom of Wagler's pit viper, Trimeresurus wagleri.";
RL Toxicon 29:227-236(1991).
CC -!- FUNCTION: Elicit tachypnea, ocular proctosis, rapid collapse
CC and spasms in mice.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -!- TISSUE SPECIFICITY: LD(50) is 0.369 mg/kg by intraperitoneal injection.
DR PIR; A39559; A39559.
KW Venom; Toxin.
FT DISULFID 9 13
FT FT PROBABLE.
SQ SEQUENCE 22 AA; 2522 MW; B0D633A32692264C CRC64;

Query Match 18.0%; Score 32.5; DB 1; Length 22;
Best Local Similarity 45.5%; Pred. No. 1.8e+02;
Matches 10; Conservative 2; Mismatches 7; Indels 3; Gaps 2;

QY 11 GITELGCGHSEPCN-THRGKP 31
| : | : | : | : | : |
DB 2 GKPDLRPCH--PPCHYIPRPKP 21
```

```
RESULT 6
MT_AGABI STANDARD; PRT; 25 AA.
ID MT_AGABI
AC P04358;
DT 20-MAR-1987 (Rel. 04, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Metallothionein (MT).
OS Agaricus bisporus (Common mushroom).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Agaricales; Agaricaceae; Agaricus.
OX NCBI_TaxID=5341;
RN [1]
RP SEQUENCE.
RC STRAIN=A-32; TISSUE=Mycelium;
RA Muenger K., Lerch K.;
RT "Copper metallothionein from the fungus Agaricus bisporus: chemical
RT and spectroscopic properties.";
RL Biochemistry 24:6751-6756(1985).
CC -!- FUNCTION: THE METALLOTHIONEINS ARE INVOLVED IN THE CELLULAR
CC SEQUESTRATION OF TOXIC METAL IONS. BINDS SIX COPPER (CUPROUS)
CC IONS.
CC -!- MISCELLANEOUS: THE ABSORPTION, LUMINESCENT, AND STEREOPTICAL
CC PROPERTIES OF THE COPPER MT ARE ATTRIBUTED TO THE METAL-THIOLATE
CC COMPLEX BECAUSE THEY ARE NOT PRESENT IN THE APOPROTEIN.
CC -!- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 8.
DR PIR; A03286; SMWR.
KW Metal-binding; Metal-thiolate cluster; Copper.
SQ SEQUENCE 25 AA; 2233 MW; 33AB70F21023CFF7 CRC64;

Query Match 17.7%; Score 32; DB 1; Length 25;
Best Local Similarity 35.7%; Pred. No. 2.5e+02;
Matches 5; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 17 GCHGSEPCNIHRGK 30
||| : | : | :
DB 4 GCSGASSTCASGQ 17

RESULT 7
CXVA_CONER STANDARD; PRT; 30 AA.
ID CXVA_CONER
AC P58782;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alpha-A-conotoxin EIVA.
OS Conus ermineus (Atlantic fish-hunting cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=55423;
RN [1]
RP SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.
RC TISSUE=Venom;
RX MEDLINE=97426393; PubMed=9278406;
RA Jacobsen R.B., Yoshikami D., Ellison M., Martinez J., Gray W.R.,
RA Cartier G.E., Shon K.-J., Groebe D.R., Abramson S.N., Olivera B.M.,
RA McIntosh J.M.;
RT "Differential targeting of nicotinic acetylcholine receptors by novel
RT alphaA-conotoxins.";
RL J. Biol. Chem. 272:22531-22537(1997).
CC -!- FUNCTION: Alpha-conotoxins act on postsynaptic membranes, they
CC bind to the nicotinic acetylcholine receptors (nAChR) and thus
CC inhibit them.
CC -!- SUBUNIT: Binds nicotinic acetylcholine receptor.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -!- MASS SPECTROMETRY: MW=3095.2; METHOD=Electrospray.
CC -!- SIMILARITY: BELONGS TO THE ALPHA-A-CONOTOXIN FAMILY.
KW Postsynaptic neurotoxin; Neurotoxin; Toxin;
KW Acetylcholine receptor inhibitor; Amidation; Hydroxylation.
```



```

OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE, AND CHARACTERIZATION.
RX MEDLINE=96019963; PubMed=7487065;
RA Biondi R.M., Veron M., Walz K., Passeron S.;
RT "Candida albicans nucleoside-diphosphate kinase: purification and
characterization.";
RL Arch. Biochem. Biophys. 323:187-194(1995).
CC -!- FUNCTION: MAJOR ROLE IN THE SYNTHESIS OF NUCLEOSIDE TRIPHOSPHATES
CC -!- CATALYTIC ACTIVITY: ATP + nucleoside diphosphate = ADP +
nucleoside triphosphate.
CC -!- SUBUNIT: HOMOHXAMER.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE NDK FAMILY.
DR HSP; P22887; ILWX.
DR InterPro; IPR001564; NDK.
DR Pfam; PF00334; NDK; 1.
DR PROSITE; PS00469; NDP_KINASES; 1.
KW Transferase; Kinase; ATP-binding.
FT NON_TER 1 1
FT ACT_SITE 4 4 BY SIMILARITY.
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2379 MW; 9DABB3A325947001 CRC64;

Query Match 16.0%; Score 29; DB 1; Length 21;
Best Local Similarity 80.0%; Pred. No. 5.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 18 CHGSE 22
Db 3 CHGSD 7

RESULT 11
GUN2_SCLSC
ID GUN2_SCLSC STANDARD; PRT; 31 AA.
AC P21834;
AT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-MAY-1991 (Rel. 18, Last annotation update)
DE Endoglucanase 2 (EC 3.2.1.4) (Endo-1,4-beta-glucanase) (Cellulase)
DE (E82) (Fragment).
OS Sclerotinia sclerotiorum.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomycetes;
OC Helotiales; Sclerotiniaceae; Sclerotinia.
OX NCBI_TaxID=5180;
RN [1]
RP SEQUENCE.
RC STRAIN=De Bary;
RX MEDLINE=91120823; PubMed=1991146;
RA Waksman G.;
RT "Purification and characterization of two endo-beta-1,4-D-glucanases
from Sclerotinia sclerotiorum.";
RL Biochim. Biophys. Acta 1073:49-55(1991).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
linkages in cellulose.
CC -!- MISCELLANEOUS: ACTIVE TOWARDS CARBOXYMETHYL CELLULOSE.
DR PIR; S13663; S13663.
KW Cellulose degradation; Hydrolase; Glycosidase.
FT NON_TER 31 31
SQ SEQUENCE 31 AA; 3190 MW; A0C6D4AC68135A68 CRC64;

Query Match 16.0%; Score 29; DB 1; Length 31;
Best Local Similarity 40.0%; Pred. No. 8.2e+02;
Matches 6; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 6 NSKFGITGELGGCH 20
Db 1 NIYYAGVAESGGEFG 15

RESULT 12

```

```

CXST_CONTE
ID CXST_CONTE STANDARD; PRT; 12 AA.
AC P58846;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Textile scratcher peptide.
OS Conus textile (Cloth-of-gold cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6494;
RN [1]
RP SEQUENCE.
RA Ramilo C.A.;
RL Thesis (1986), University of Manila, Philippines.
CC -!- FUNCTION: Causes scratching in mice.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -!- SIMILARITY: BELONGS TO THE M-SUPERFAMILY OF CONOTOXINS.
KW Toxin; Amidation; Hydroxylation.
FT DISULFID 1 8 PROBABLE.
FT DISULFID 2 11 PROBABLE.
FT DISULFID 5 12 PROBABLE.
FT MOD_RES 10 10 HYDROXYLATION.
FT MOD_RES 12 12 AMIDATION.
SQ SEQUENCE 12 AA; 1297 MW; 81FD6FAC21B879CE CRC64;

Query Match 15.5%; Score 28; DB 1; Length 12;
Best Local Similarity 57.1%; Pred. No. 4.7e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 18 CHGSEPC 24
Db 5 CFGCTPC 11

RESULT 13
VSPL_AKGBI
ID VSPL_AKGBI STANDARD; PRT; 20 AA.
AC P33588;
AT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 40, Last annotation update)
DE Anctod (EC 3.4.21.74) (Venombin A) (Protein C activator) (ACC-C)
DE (Fragment).
OS Agkistrodon bilineatus (Cantill) (Tropical moccasin).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Agkistrodon.
OX NCBI_TaxID=8718;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=90350102; PubMed=2385829;
RA Nakagaki T., Kazim A.L., Kisiel W.;
RT "Isolation and characterization of a protein C activator from
tropical moccasin venom.";
RL Thromb. Res. 58:593-602(1990).
CC -!- FUNCTION: THROMBIN-LIKE SNAKE VENOM SERINE PROTEASE. ACTIVATES
PROTEIN C.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Arg-|-Xaa bond in
fibrinogen, to form fibrin, and release fibrinopeptide A. The
specificity of further degradation of fibrinogen varies with
species origin of the enzyme.
CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC -!- TISSUE SPECIFICITY: EXPRESSED BY VENOM GLAND.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. SNAKE VENOM SUBFAMILY.
DR PIR; A60489; A60489.
DR MEROPS; S01.178; -.
DR InterPro; IPR001254; Ser_protease_Try.
DR PROSITE; PS0240; TRYPSIN_DOM; PARTIAL.
DR PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
DR PROSITE; PS00135; TRYPSIN_SER; PARTIAL.

```

KW Hydrolase; Serine protease; Venom; Glycoprotein.  
 FT NON\_TER 20  
 SQ SEQUENCE 20 AA; 2191 MW; 6E99F8B4CC53EFE1 CRC64;

Query Match 15.5%; Score 28; DB 1; Length 20;  
 Best Local Similarity 54.5%; Pred. No. 7.6e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

Oy 20 GSEPCNI--HR 28  
 | : ||| ||  
 Db 3 GGDECNINEHR 13

## RESULT 14

VSPI\_TRIOK STANDARD; PRT; 20 AA.  
 ID VSP1\_TRIOK  
 AC P20005;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Okinaxobin I (EC 3.4.21.-) (Fragment).  
 OS Trimeresurus okinavensis (Hime-habu) (Ovophis okinavensis).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;  
 OC Viperidae; Crotalinae; Ovophis.  
 OX NCBI\_TaxID=8769;  
 RN [1]  
 RP SEQUENCE.

TISSUE-Venom;  
 RX MEDLINE=91185341; PubMed=19644457;  
 RA Iwasaki A., Shieh T.-C., Shimohigashi Y., Waki M., Kihara H., Ohno M.;  
 RT "Purification and characterization of a coagulant enzyme, okinaxobin  
 I, from the venom of Trimeresurus okinavensis (Himehabu snake) which  
 releases fibrinopeptide B.";  
 RL J. Biochem. 108:822-828(1990).

CC -!- FUNCTION: THROMBIN-LIKE SNAKE VENOM SERINE PROTEASE.  
 CC -!- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-GLY BONDS IN  
 CC FIBRINOGEN ALPHA CHAINS. CLEAVES FIBRINOGEN ALPHA CHAINS  
 CC PREFERENTIALLY TO BETA CHAINS.  
 CC -!- SUBUNIT: MONOMER.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: EXPRESSED BY VENOM GLAND.  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. SNAKE VENOM SUBFAMILY.

PIR; PX0042; PX0042.  
 DR MEROPS; S01.185; -;  
 DR InterPro: IPR001254; Ser\_protease\_Try.

DR PROSITE; PS50240; TRYPSIN\_DOM; PARTIAL.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; PARTIAL.  
 DR PROSITE; PS00135; TRYPSIN\_SER; PARTIAL.

KW Hydrolase; Serine protease; Venom.  
 FT NON\_TER 1  
 SQ SEQUENCE 20 AA; 2249 MW; 69E9EF62EC53F391 CRC64;

Query Match 15.5%; Score 28; DB 1; Length 20;  
 Best Local Similarity 54.5%; Pred. No. 7.6e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

Oy 20 GSEPCNI--HR 28  
 | : ||| ||  
 Db 3 GGDECNINEHR 13

## RESULT 15

NRLA\_ACISP STANDARD; PRT; 21 AA.  
 ID NRLA\_ACISP  
 AC P33036;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Nitrilase (EC 3.5.5.1) (Fragment).  
 OS Acinetobacter sp. (strain AK226).  
 OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;  
 OC Acinetobacter.

OX NCBI\_TaxID=472;  
 RN [1]  
 RP SEQUENCE.

RX MEDLINE=91345837; PubMed=1369128;  
 RA Yamamoto K., Komatsu K.;  
 RT "Purification and characterization of nitrilase responsible for the  
 enantioselective hydrolysis from Acinetobacter sp. AK 226.";  
 RL Agric. Biol. Chem. 55:1459-1466(1991).  
 CC -!- FUNCTION: ACTS ON MANY KINDS OF NITRILE COMPOUNDS SUCH AS  
 CC ALIPHATIC, AROMATIC, AND HETEROCYCLIC MONONITRILES OR DINITRILES.  
 CC PREFERS S-(-)-2-(4'-ISOBUTYLPHENYL)-PROPIONITRILE TO R-(+)-2-(4'-  
 CC ISOBUTYLPHENYL)-PROPIONITRILE AS THE SUBSTRATE.  
 CC -!- CATALYTIC ACTIVITY: A nitrile + H(2)O = a carboxylate + NH(3).  
 CC -!- SIMILARITY: BELONGS TO THE NITRILASE FAMILY.

DR PIR; JU0386; JU0386.

DR InterPro: IPR000132; N/CN\_hydrolase.

DR PROSITE; PS00920; NITRIL\_CHT\_1; PARTIAL.

DR PROSITE; PS00921; NITRIL\_CHT\_2; PARTIAL.

KW Hydrolase.

FT NON\_TER 21 21

SQ SEQUENCE 21 AA; 2223 MW; 5FA741C41EAC619B CRC64;

Query Match 15.5%; Score 28; DB 1; Length 21;

Best Local Similarity 45.5%; Pred. No. 7.9e+02;

Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Oy 3 IRANSKFIGIT 13

: ||||: |

Db 1 VSYNSKFLAAT 11

Search completed: December 4, 2002, 13:17:08  
 Job time : 11 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 13:15:37 ; Search time 29 Seconds  
(without alignments)  
227.362 Million cell updates/sec

Title: US-09-362-731A-5  
 Perfect score: 181  
 Sequence: 1 QYIKANSKFTGITELGCGHSGSEPCNTHRGKPF 32

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 16362

```
Minimum DB seq length: 0
Maximum DB seq length: 32
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

```
Database :
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*
```

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	23.2	27	4	Q8tb35 homo sapien
2	40	22.1	25	2	Q9zew1 enterobacte
3	37	20.4	30	2	Q93gf6 staphylococ
4	35	19.3	27	13	Q9dfb4 brachydanio
5	34	18.8	29	13	Q91396 gallus gall
6	33.5	18.5	22	5	Q9ndt8 balanus amp
7	33.5	18.5	22	5	Q9ndS9 balanus amp
8	33.5	18.5	23	7	Q95376 homo sapien
9	33.5	18.5	31	12	Q68018 hepatitis b
10	33.5	18.5	31	12	Q68054 hepatitis b
11	33.5	18.5	31	12	Q68080 hepatitis b
12	33	18.2	23	6	Q9tre0 ovis aries
13	32.5	18.0	28	12	Q68009 hepatitis b
14	32.5	18.0	31	12	Q67957 hepatitis b
15	32.5	18.0	31	12	Q67960 hepatitis b
16	32.5	18.0	31	12	Q67962 hepatitis b

17	32.5	18.0	31	12	Q67965	Q67965	hepatitis b
18	32.5	18.0	31	12	Q67967	Q67967	hepatitis b
19	32.5	18.0	31	12	Q67969	Q67969	hepatitis b
20	32.5	18.0	31	12	Q67971	Q67971	hepatitis b
21	32.5	18.0	31	12	Q67977	Q67977	hepatitis b
22	32.5	18.0	31	12	Q67979	Q67979	hepatitis b
23	32.5	18.0	31	12	Q67981	Q67981	hepatitis b
24	32.5	18.0	31	12	Q67983	Q67983	hepatitis b
25	32.5	18.0	31	12	Q67985	Q67985	hepatitis b
26	32.5	18.0	31	12	Q67987	Q67987	hepatitis b
27	32.5	18.0	31	12	Q67990	Q67990	hepatitis b
28	32.5	18.0	31	12	Q67992	Q67992	hepatitis b
29	32.5	18.0	31	12	Q67996	Q67996	hepatitis b
30	32.5	18.0	31	12	Q67998	Q67998	hepatitis b
31	32.5	18.0	31	12	Q68000	Q68000	hepatitis b
32	32.5	18.0	31	12	Q68002	Q68002	hepatitis b
33	32.5	18.0	31	12	Q68007	Q68007	hepatitis b
34	32.5	18.0	31	12	Q68013	Q68013	hepatitis b
35	32.5	18.0	31	12	Q68015	Q68015	hepatitis b
36	32.5	18.0	31	12	Q68017	Q68017	hepatitis b
37	32.5	18.0	31	12	Q68019	Q68019	hepatitis b
38	32.5	18.0	31	12	Q68021	Q68021	hepatitis b
39	32.5	18.0	31	12	Q68024	Q68024	hepatitis b
40	32.5	18.0	31	12	Q68026	Q68026	hepatitis b
41	32.5	18.0	31	12	Q68028	Q68028	hepatitis b
42	32.5	18.0	31	12	Q68029	Q68029	hepatitis b
43	32.5	18.0	31	12	Q68031	Q68031	hepatitis b
44	32.5	18.0	31	12	Q68033	Q68033	hepatitis b
45	32.5	18.0	31	12	Q68034	Q68034	hepatitis b

## ALIGNMENTS

## RESULT 1

Q8TB35  
ID Q8TB35 PRELIMINARY; PRT: 27 AA.

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OS Homo sapiens (Human).

RESULT 2  
Q9ZEW1  
ID Q9ZEW1  
PRELIMINARY;  
PRT: 25 AA.

GN TAPA.  
OG Enterobacter intermedium.  
OS Plasmid pLV1402.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Enterobacter.  
OX NCBI\_TaxID=61648;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=AH14;  
RA MEDLINE=20431899; PubMed=10974114;  
RA Osborn A.M., Da Silva Tatley F.M., Steyn L.M., Pickup R.W.,  
RT Saunders J.R.;  
RT "Mosaic plasmids and mosaic replicons: evolutionary lessons from the  
RT analysis of genetic diversity in IncFII-related replicons.";  
RL Microbiology 146:2267-2275(2000).  
DR EMBL: AJ009981; CAA08929.1; -  
KW Plasmid.  
SQ SEQUENCE 25 AA; 2922 MW; A904596B6082C837 CRC64;  
  
Query Match 22.1%; Score 40; DB 2; Length 25;  
Best Local Similarity 53.8%; Pred. No. 49;  
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
  
QY 18 CHGSEPCNIHRGK 30  
II IIII I:  
Db 11 CHLLLPNISAGR 23  
  
RESULT 3  
Q93GF6 PRELIMINARY; PRT; 30 AA.  
AC Q93GF6;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Aurb.  
GN Aurb.  
OS Staphylococcus aureus.  
OG Plasmid pR36.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
OC Staphylococcus.  
OX NCBI\_TaxID=1280;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21424233; PubMed=11531330;  
RA Netz D.J.A., Sahl H.-G., Marcolino R., Nascimento J.S., Oliveira S.S.,  
RA Soares M.B., Bastos M.C.F.;  
RT "Molecular characterisation of aureocin A70, a multi-peptide  
RT bacteriocin isolated from Staphylococcus aureus.";  
RL J. Mol. Biol. 311:939-949(2001).  
DR EMBL: AF241888; AAK73553.1; -  
KW Plasmid.  
SQ SEQUENCE 30 AA; 2797 MW; 4B60654093F92DFA CRC64;  
  
Query Match 20.4%; Score 37; DB 2; Length 30;  
Best Local Similarity 50.0%; Pred. No. 1.7e+02;  
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
  
QY 5 ANSKFIGITELGGCHG 20  
I :II:I IIII I  
Db 3 AVAKFLGKALGGAAG 18  
  
RESULT 4  
Q9DFB4 PRELIMINARY; PRT; 27 AA.  
AC Q9DFB4;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE z-2-F1 alpha isoform 2 (Fragment).  
GN NR5A2 OR FFLA.  
OS Brachydanio rerio (Zebrafish) (Zebra danio).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20277948; PubMed=10816440;  
RA Lin W., Wang H.W., Sum C., Liu D., Hew C.L., Chung B.;  
RT "zebrafish ftz-f1 gene has two promoters, is alternatively spliced,  
RT and is expressed in digestive organs.";  
RL Biochem. J. 348:439-446(2000).  
DR EMBL: AF286364; AAG28163.1; -  
DR ZFIN; ZDB-GENE-990415-79; nr5a2.  
FT NON\_TER 27 27  
SQ SEQUENCE 27 AA; 2997 MW; 2783518B2738CEC9 CRC64;  
  
Query Match 19.3%; Score 35; DB 13; Length 27;  
Best Local Similarity 30.4%; Pred. No. 3.1e+02;  
Matches 7; Conservative 5; Mismatches 11; Indels 0; Gaps 0;  
  
QY 4 KANSKEIGITELGGCHGSEPCNI 26  
I :II:I I:  
Db 4 KVESEYLGARSHGEGQHMPGNM 26  
  
RESULT 5  
Q91396 PRELIMINARY; PRT; 29 AA.  
AC Q91396;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE Anion exchanger AE1-5 protein (Fragment).  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=KIDNEY;  
RX MEDLINE=95208819; PubMed=7900851;  
RA Cox K.H., Cox J.V.;  
RT "Variant chicken AE1 anion exchangers possess divergent NH(2)-terminal  
RT cytoplasmic domains.";  
RL Am. J. Physiol. 268:F503-F513(1995).  
DR EMBL: S76801; AAD14986.2; -  
FT NON\_TER 1 1  
FT NON\_TER 29 29  
SQ SEQUENCE 29 AA; 2904 MW; 71280BFCE0743095 CRC64;  
  
Query Match 18.8%; Score 34; DB 13; Length 29;  
Best Local Similarity 43.8%; Pred. No. 4.8e+02;  
Matches 7; Conservative 1; Mismatches 4; Indels 4; Gaps 1;  
  
QY 16 GGCHGSEPCNIHRGKP 31  
III I :II I  
Db 2 GGCVG----GVHRADP 13  
  
RESULT 6  
Q9NDT8 PRELIMINARY; PRT; 22 AA.  
AC Q9NDT8;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE BCS-6.  
GN BCS-6.  
OS Balanus amphitrite (Barnacle).  
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Maxillopoda; Cirripedia;  
OC Thoracica; Sessilia; Balanomorpha; Balanoidae; Balanus.  
OX NCBI\_TaxID=32267;

DR	EMBL; U50061; AAB63532.1; -
FT	NON_TER 1
SQ	SEQUENCE 23 AA; 2729 MW; AD6B35DC3A259EBF CRC64;
 Query Match 18.5%; Score 33.5; DB 7; Length 23; Best Local Similarity 53.8%; Pred. No. 4.5e+02; Matches 7; Conservative 1; Mismatches 2; Indels 3; Gaps 1;	
Qy	18 CH--GSEPCNIH 27     :
Dd	10 CHFLNGTEPDQIH 22     :
 RESULT 9	
Q68018	
ID	Q68018 PRELIMINARY; PRT; 31 AA.
AC	Q68018:
DT	01-NOV-1996 (TrEMBLrel. 01, Created)
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE	x, PreC and C genes (Castag) (Fragment).
GN	X.
OS	Hepatitis B virus.
OC	Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX	NCBI_TaxID=10407;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Lai M.E., Mazzoleni A.P., Porru A., Balestrieri A.;
RL	Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
DR	EMBL; X85290; CAAS9608.1; -
DR	InterPro; IPR000236; TransactX.
PFam	Pfam; PF00739; X; 1.
FT	NON_TER 1
SQ	SEQUENCE 31 AA; 3344 MW; 0873FB70509B8D79 CRC64;
 Query Match 18.5%; Score 33.5; DB 12; Length 31; Best Local Similarity 38.1%; Pred. No. 6.2e+02; Matches 8; Conservative 1; Mismatches 7; Indels 5; Gaps 1;	
Qy	10 IGITELGGCH----GSEPCN 25   :
Dd	6 KVIILGGCRHKLVCAAPCN 26   :
 RESULT 10	
Q68054	
ID	Q68054 PRELIMINARY; PRT; 31 AA.
AC	Q68054;
DT	01-NOV-1996 (TrEMBLrel. 01, Created)
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE	x, PreC and C genes (Mannoni 1) (Fragment).
GN	X.
OS	Hepatitis B virus.
OC	Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX	NCBI_TaxID=10407;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Lai M.E., Mazzoleni A.P., Porru A., Balestrieri A.;
RL	Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
DR	EMBL; X85309; CAAS9649.1; -
DR	InterPro; IPR000236; TransactX.
PFam	Pfam; PF00739; X; 1.
FT	NON_TER 1
SQ	SEQUENCE 31 AA; 3346 MW; BLACAB2BB08DEA72 CRC64;
 Query Match 18.5%; Score 33.5; DB 12; Length 31; Best Local Similarity 43.8%; Pred. No. 6.2e+02; Matches 7; Conservative 0; Mismatches 4; Indels 5; Gaps 1;	
Qy	15 LGGCH-----GSEPCN 25 
Dd	11 LGGCRHKLVCAAPCN 26 



## RESULT 11

Q68080 PRELIMINARY; PRT; 31 AA.  
 ID Q68080;  
 AC Q68080;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
 DE X, PreC and C genes (Pische) (Fragment).  
 GN X.  
 OS Hepatitis B virus.  
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
 OX NCBI\_TaxID=10407;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lai M.E., Mazzoleni A.P., Porru A., Balestrieri A.;  
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; X85262; CAA59534.1; -.  
 DR InterPro: IPR000236; TransactX.  
 DR Pfam; PF00739; X; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 31 AA; 3394 MW; E043AB25508DEA64 CRC64;

Query Match 18.5%; Score 33.5; DB 12; Length 31;  
 Best Local Similarity 43.8%; Pred. No. 6.2e+02;  
 Matches 7; Conservative 0; Mismatches 4; Indels 5; Gaps 1;  
 QY 15 LGGCH-----GSEPCN 25  
 |||||  
 Db 11 LGGCRHKLVCAPPCN 26

## RESULT 12

Q9TRE0 PRELIMINARY; PRT; 23 AA.  
 ID Q9TRE0;  
 AC Q9TRE0;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
 DE 57 kDa PLACENTATION-SPECIFIC BINUCLEATE cell glycoprotein.  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoides;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=94075364; PubMed=8253801;  
 RA Atkinson Y.H., Gogolin-Ewens K.J., Hounsell E.F., Davies M.J.,  
 RA Brandon M.R., Seemark R.F.;  
 RT "Characterization of placental-specific binucleate cell  
 RT glycoproteins possessing a novel carbohydrate. Evidence for a new  
 RT family of pregnancy-associated molecules.";  
 RL J. Biol. Chem. 268:26679-26685(1993).  
 SQ SEQUENCE 23 AA; 2721 MW; 1818555496F40F8C CRC64;

Query Match 18.2%; Score 33; DB 6; Length 23;  
 Best Local Similarity 62.5%; Pred. No. 5.3e+02;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 25 NTHRGKPF 32  
 |||||

Db 15 NTHXGNPY 22

## RESULT 13

Q68009 PRELIMINARY; PRT; 28 AA.  
 ID Q68009;  
 AC Q68009;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last annotation update)

DE X, PreC and C genes (Ferracuti 2) (Fragment).  
 GN X.  
 OS Hepatitis B virus.  
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
 OX NCBI\_TaxID=10407;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lai M.E., Mazzoleni A.P., Porru A., Balestrieri A.;  
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; X85285; CAA59597.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 28 AA; 2982 MW; 8D7D0FC39676E80D CRC64;

Query Match 18.0%; Score 32.5; DB 12; Length 28;  
 Best Local Similarity 43.8%; Pred. No. 7.9e+02;  
 Matches 7; Conservative 0; Mismatches 4; Indels 5; Gaps 1;

QY 15 LGGCH-----GSEPCN 25  
 |||||  
 Db 8 LGGCRHKLVCAPPCN 23

## RESULT 14

Q67957 PRELIMINARY; PRT; 31 AA.  
 ID Q67957;  
 AC Q67957;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
 DE X, PreC and C genes (Tarisloc) (Fragment).  
 GN X.  
 OS Hepatitis B virus.  
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
 OX NCBI\_TaxID=10407;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lai M.E., Mazzoleni A.P., Porru A., Balestrieri A.;  
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; X85255; CAA59515.1; -.  
 DR InterPro: IPR000236; TransactX.  
 DR Pfam; PF00739; X; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 31 AA; 3350 MW; A776AB3BA162EA72 CRC64;

Query Match 18.0%; Score 32.5; DB 12; Length 31;  
 Best Local Similarity 43.8%; Pred. No. 8.8e+02;  
 Matches 7; Conservative 0; Mismatches 4; Indels 5; Gaps 1;

QY 15 LGGCH-----GSEPCN 25  
 |||||  
 Db 11 LGGCRHKLVCAPPCN 26

Query Match 18.0%; Score 32.5; DB 12; Length 31;  
 Best Local Similarity 43.8%; Pred. No. 8.8e+02;  
 Matches 7; Conservative 0; Mismatches 4; Indels 5; Gaps 1;

QY 15 LGGCH-----GSEPCN 25  
 |||||  
 Db 11 LGGCRHKLVCAPPCN 26

## RESULT 15

Q67960 PRELIMINARY; PRT; 31 AA.  
 ID Q67960;  
 AC Q67960;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
 DE X, PreC and C genes (Buesca) (Fragment).  
 GN X.  
 OS Hepatitis B virus.  
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
 OX NCBI\_TaxID=10407;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lai M.E., Mazzoleni A.P., Porru A., Balestrieri A.;  
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; X85264; CAA59539.1; -.  
 DR InterPro: IPR000236; TransactX.  
 DR Pfam; PF00739; X; 1.  
 FT NON\_TER 1

Query Match 18.0%; Score 32.5; DB 12; Length 31;  
 Best Local Similarity 43.8%; Pred. No. 8.8e+02;  
 Matches 7; Conservative 0; Mismatches 4; Indels 5; Gaps 1;

QY 15 LGGCH-----GSEPCN 25  
 |||||  
 Db 11 LGGCRHKLVCAPPCN 26

Query Match 18.0%; Score 32.5; DB 12; Length 31;  
 Best Local Similarity 43.8%; Pred. No. 8.8e+02;  
 Matches 7; Conservative 0; Mismatches 4; Indels 5; Gaps 1;

SQ SEQUENCE 31 AA; 3397 MW; 5FF9CD25508DFC08 CRC64;  
Query Match 18.0%; Score 32.5; DB 12; Length 31;  
Best Local Similarity 43.8%; Pred. No. 8.8e+02;  
Matches 7; Conservative 0; Mismatches 4; Indels 5; Gaps 1;  
QY 15 LGGCH-----GSEPCN 25  
Db 11 LGGCRHKLVCAPAPCN 26  
|||  
|||

Search completed: December 4, 2002, 13:17:44  
Job time : 30 secs



```

DR WPI; 2000-422470/36.
XX
PT New compound for prevention and treatment of allergies comprises at
PT least one allergen antigenic determinant recognized by a B cell and at
PT least one antigenic determinant which does not trigger T cell
PT activation -
XX
PS Claim 8; Page 35; 50pp; English.
XX
CC The present invention describes a compound (I) for the prevention and/or
CC treatment of allergy. The compound comprises at least one allergen
CC antigenic determinant (i) recognised by a B cell or an antibody secreted
CC by a B cell of a non-atopic individual and at least one antigenic
CC determinant (ii) different from the allergen that triggers T cell
CC activation. (I) has antiallergic, antiasthmatic, antiinflammatory,
CC dermatological and immunosuppressive activities, and can be used in a
CC vaccine. (I) may be used in a pharmaceutical or cosmetic medicament to
CC treat and/or prevent allergies or a disease of allergic origin,
CC especially hypersensitivities. These include rhinitis, sinusitis,
CC bronchial asthma, atopic dermatitis, some forms of acute and chronic
CC urticaria, gastro-intestinal syndromes associated with the ingestion of
CC food allergens, oro-pharyngeal syndrome, anaphylactic reactions
CC associated with drug hypersensitivities and/or a mixture of these. The
CC use of (I) in the treatment of allergic conditions avoids the need for
CC drug treatment, which often causes undesirable side-effects. Also, prior
CC art drug therapies alleviate symptoms, but do not influence their
CC causes, however (I) actually combats the cause of an allergic reaction.
CC The present sequence represents a specifically claimed compound peptide
CC sequence from the present invention.
XX
SQ Sequence 32 AA;
Query Match 100.0%; Score 181; DB 21; Length 32;
Best Local Similarity 100.0%; Pred. No. 6e-19;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QYIKANSKFIGITELGGCHGSEPCNIHRGKPF 32
Db 1 QYIKANSKFIGITELGGCHGSEPCNIHRGKPF 32
RESULT 2
ID AAY82632 standard; peptide; 31 AA.
XX
AC AAY82632;
XX
DT 07-AUG-2000 (first entry)
XX
DE Tetanus toxoid T cell epitope and Der pII B cell epitope peptide.
XX
KW T cell epitope; B cell epitope; allergy; allergen; antigenic;
KW antiallergic; antiasthmatic; antiinflammatory; dermatological;
KW immunosuppressive; vaccine; rhinitis; sinusitis; bronchial asthma;
KW atopic dermatitis; acute urticaria; chronic urticaria;
KW gastro-intestinal syndrome; food allergen; oro-pharyngeal syndrome;
KW anaphylactic reaction; drug hypersensitivity; allergic reaction.
XX
OS Dermatophagoides pteronyssinus.
OS Clostridium tetani.
OS Synthetic.
XX
PN WO200006694-A2.
XX
PD 10-FEB-2000.
XX
PF 20-JUL-1999; 99WO-BE00092.
XX
PR 30-JUL-1998; 98EP-0870167.
XX
PA (UNTO) ÜCB SA.
XX
PI Saint-Remy J, Jacquemin M;

```

---

```

XX WPI; 2000-422470/36.
XX
PT New compound for prevention and treatment of allergies comprises at
PT least one allergen antigenic determinant recognized by a B cell and at
PT least one antigenic determinant which does not trigger T cell
PT activation -
XX
PS Claim 8; Page 35; 50pp; English.
XX
CC The present invention describes a compound (I) for the prevention and/or
CC treatment of allergy. The compound comprises at least one allergen
CC antigenic determinant (i) recognised by a B cell or an antibody secreted
CC by a B cell of a non-atopic individual and at least one antigenic
CC determinant (ii) different from the allergen that triggers T cell
CC activation. (I) has antiallergic, antiasthmatic, antiinflammatory,
CC dermatological and immunosuppressive activities, and can be used in a
CC vaccine. (I) may be used in a pharmaceutical or cosmetic medicament to
CC treat and/or prevent allergies or a disease of allergic origin,
CC especially hypersensitivities. These include rhinitis, sinusitis,
CC bronchial asthma, atopic dermatitis, some forms of acute and chronic
CC urticaria, gastro-intestinal syndromes associated with the ingestion of
CC food allergens, oro-pharyngeal syndrome, anaphylactic reactions
CC associated with drug hypersensitivities and/or a mixture of these. The
CC use of (I) in the treatment of allergic conditions avoids the need for
CC drug treatment, which often causes undesirable side-effects. Also, prior
CC art drug therapies alleviate symptoms, but do not influence their
CC causes, however (I) actually combats the cause of an allergic reaction.
CC The present sequence represents a specifically claimed compound peptide
CC sequence from the present invention.
XX
SQ Sequence 31 AA;
Query Match 54.1%; Score 98; DB 21; Length 31;
Best Local Similarity 67.7%; Pred. No. 5.3e-07;
Matches 21; Conservative 0; Mismatches 0; Indels 10; Gaps 1;
QY 1 QYIKANSKFIGITELG-----GCHGS 21
Db 1 QYIKANSKFIGITELGGHEIKVLVPCGHS 31
RESULT 3
ID AAR51764 standard; Protein; 24 AA.
XX
AC AAR51764;
XX
DT 01-FEB-1995 (first entry)
XX
DE Der p II derived peptide, DP II-2(11-35).
XX
KW Group I; protein allergen; house dust mite; D. pteronyssinus; Der p I;
KW homology; D. farinae; Der f I; group II; Der p II; Der f II; T-cell;
KW epitopes; fusion peptides; antigenic fragments; substitution; deletion;
KW addition; chemical synthesis; chemical cleavage; recombinant techniques;
KW allergic response; immunoglobulin E; IgE; immunotherapy; anaphylaxis;
KW IgE-mediated responses; anergise; lymphokine secretion profile; modify;
KW T cell subpopulations; unresponsive; Immune response; tolerance.
XX
OS Dermatophagoides pteronyssinus.
XX
PN ZA9302677-A.
XX
PD 26-JAN-1994.
XX
PF 16-APR-1993; 93ZA-0002677.
XX
PR 16-APR-1993; 93ZA-0002677.
XX
PA (IMMU-) IMMULOGIC PHARM CORP.
XX
PI Garman RD, Greenstein JL, Kuo M, Rogers BL;

```

XX WPI; 1994-126807/15.

XX Isolated and/or modified peptides comprising T-cell epitopes - of

PT major protein allergens of genus Dermatophagoides, used to treat

PT or diagnose sensitivity to house dust mites

XX Disclosure; Fig 3; 154pp; English.

XX The sequences given in AAR51731-841 represent T-cell epitopes derived

CC from the group I and II protein allergens from the house dust mite D.

CC farinae and D. pteronyssinus, Der f I, Der f II, Der p I and Der p II

CC respectively. The Der f II proteinsCC shows high homology having an

CC identity of 88%, with an identity of 81% between the two group I

CC proteins (see also AAR51727-30). Fusion peptides may be produced which

CC comprise at least two or these antigenic fragments. Each region of

CC these fusion peptides may be derived from the same, or different, mite

CC allergens. The antigenic fragments may be altered by substitution,

CC deletion or addition to enhance their antigenicity. These peptides may

CC be produced by chemical synthesis, chemical cleavage of the protein

CC allergen or by recombinant techniques. These peptides, or the fusion

CC peptides, when administered to a house dust mite sensitive individual,

CC are capable of modifying the allergic response of the individual to the

CC allergen. The peptides do not bind to immunoglobulin E (IgE), or bind

CC IgE to a lesser extent than the full length protein allergen. This

CC reduces the major complications of standard immunotherapy, which are

CC IgE-mediated responses such as anaphylaxis. Exposure of mite allergic

CC patients to these peptides may tolerate or anergise appropriate T cell

CC subpopulations such that they become unresponsive to mite allergens and

CC do not participate in mounting an immune response upon exposure.

CC Administration of the peptides may also modify the lymphokine secretion

CC profile as compared with exposure to the naturally occurring mite protein

CC allergen.

XX

SQ Sequence 24 AA;

Query Match 50.8%; Score 92; DB 15; Length 24;

Best Local Similarity 93.8%; Pred. No. 3e-06;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 17 GCGHSEPCNIHRGKPF 32

Db | | | | | | | | | | | | | | | |

9 GCGHSEPCNIHRGKPF 24

RESULT 4

AAAR36416

ID AAR36416 standard; peptide; 25 AA.

XX

AC AAR36416;

XX

DT 12-AUG-1993 (first entry)

XX

DE DpII-2(11-35) a Dermatophagoides protein allergen.

XX

KW T cell epitope; house dust mite; allergy; soluble; Der pII.

XX

OS Synthetic.

XX

PN WO9308279-A.

XX

PD 29-APR-1993.

XX

XX 15-OCT-1992; 92WO-US08637.

PF

PR 16-OCT-1991; 91US-0777859.

PR

XX 08-MAY-1992; 92US-0881396.

XX

PA (IMMU-) IMMULOGIC PHARM CORP.

XX

XX Garman RD, Greenstein JL, Kuo MC, Rogers BL;

PI WPI; 1993-152472/18.

DR

XX Isolated peptide(s) of Dermatophagoides protein allergens - for

PT diagnosis and treatment of sensitivity to house dust mite

XX Claim 44; Fig 3; 176pp; English.

XX

CC The peptide is one of a series of overlapping peptides synthesised by

CC standard techniques to cover the whole Dermatophagoides

CC pteronyssinus der pII sequence. The T cell epitopes of the protein

CC were mapped by detection of the peptide's ability to stimulate T

CC cell activity. The peptides may be used for diagnosis and treatment

CC of sensitivity to house dust mite allergens, when administered to

CC house dust mite sensitive individuals, the peptides are capable of

CC modifying the allergic response to the allergens. The peptides may

CC be modified for e.g. increasing solubility, enhancing therapeutic or

CC preventive efficacy or stability.

CC See also AAR34686-700 and AAR36398-490.

XX

SQ Sequence 25 AA;

Query Match 50.8%; Score 92; DB 14; Length 25;

Best Local Similarity 93.8%; Pred. No. 3.1e-06;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 17 GCGHSEPCNIHRGKPF 32

Db | | | | | | | | | | | | | | | |

10 GCGHSEPCNIHRGKPF 25

RESULT 5

AAAR82550

ID AAR82550 standard; peptide; 25 AA.

XX

AC AAR82550;

XX

DT 30-MAY-1996 (first entry)

XX

DE Dermatophagoides pteronyssinus group II peptide pPII-1 (11-35).

XX

KW House dust mite; DerpI; DerfII; DerpII; allergen; allergy;

XX overlapping peptide; screening.

XX

OS Dermatophagoides pteronyssinus.

XX

PN WO9528424-A1.

XX

PD 26-OCT-1995.

XX

PF 12-APR-1995; 95WO-US04481.

XX

PR 14-APR-1994; 94US-0227722.

XX

PA (IMMU-) IMMULOGIC PHARM CORP.

XX

PI Chen X, Evans S, Franzen HM, Kuo M, Shaked Z;

XX WPI; 1995-373765/48.

DR

XX Compens. contg. house mite allergen-derived peptide(s), some of

PT which are new - are used to treat allergy, and are stable, soluble

PT and able to induce T cell non-responsiveness

XX

PS Disclosure; Fig 2; 61pp; English.

XX

CC Claimed therapeutic compositions contain at least one of the peptides

CC DPI-21.2 and DPI-22.2 and also at least one of the new peptides

CC DPI-23.31, DPI-26.6, DPII-20.9, DPII-22.14 and DPII-25.15. The

CC compositions are useful for treating sensitivity to house dust mite

CC allergens. The peptides were identified by screening overlapping

CC peptides derived from D.pteronyssinus and D.farinae group I and II

CC allergens for T-cell reactivity in sensitised individuals.

CC The present sequence is that of overlapping peptide DPII-2 (11-35).

XX

SQ Sequence 25 AA;  
 Query Match 50.8%; Score 92; DB 16; Length 25;  
 Best Local Similarity 93.8%; Pred. No. 3.1e-06;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 17 GCHGSEPCNIHRGKPF 32  
 DB 10 GCHGSEPCIIHRGKPF 25  
 ||||| |||||

RESULT 6  
 AAW71939  
 ID AAW71939 standard; peptide; 25 AA.  
 XX AC AAW71939;  
 XX DT 16-DEC-1998 (first entry)  
 XX DE Dermatophagoides sp.  
 XX KW Allergen; house dust mite; detection; sensitivity; T cell epitope;  
 XX KW screening; allergic disorder; asthma; rhinitis; ectopic dermatitis;  
 XX KW Der f I; Der p II; Der f II; house dust mite allergy.  
 XX OS Dermatophagoides sp.  
 XX PN US5820862-A.  
 XX PD 13-OCT-1998.  
 XX PF 07-JUN-1995; 95US-0482142.  
 XX PR 19-MAY-1995; 95US-0445307.  
 XX PR 14-APR-1994; 94US-0227772.  
 XX PR 07-JUN-1995; 95US-0482142.  
 XX PA (IMMU-) IMMULOGIC PHARM CORP.  
 XX PI Chen X, Evans S, Franzen HM, Garman RD, Greenstein JL;  
 XX PI Kuo M, Rogers BL, Shaked Z;  
 XX WPI: 1998-567590/48.  
 XX DT Dermatophagoides allergen peptides - useful for treating house dust  
 XX PT mite allergy  
 XX PS Disclosure: Column 99-100; 155pp; English.  
 XX CC The present invention describes peptides for treating sensitivity to  
 CC house dust mite allergens from the genus Dermatophagoides. Peptides  
 CC within the scope of the invention comprise at least one T cell epitope,  
 CC or preferably at least two T cell epitopes of a protein allergen  
 CC selected from the allergens Der p I, Der p II, Der f I, or Der f II.  
 CC The invention also describes modified peptides having similar or  
 CC enhanced therapeutic properties as the corresponding, naturally  
 CC occurring allergen, but having reduced side effects. AAW71912 to  
 CC AAW72000, and AAW72257 to AAW72330 represent peptides from the present  
 CC invention.

SQ Sequence 25 AA;  
 Query Match 50.8%; Score 92; DB 19; Length 25;  
 Best Local Similarity 93.8%; Pred. No. 3.1e-06;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 17 GCHGSEPCNIHRGKPF 32  
 DB 10 GCHGSEPCIIHRGKPF 25  
 ||||| |||||

RESULT 7  
 AAY50393  
 ID AAY50393 standard; Peptide; 25 AA.

XX AC AAY50393;  
 XX DT 25-JAN-2000 (first entry)  
 XX DE Dermatophagoides sp major protein allergen DP II-2.  
 XX KW Allergen; house dust mite; detection; sensitivity; T cell epitope;  
 XX KW screening; allergic disorder; asthma; rhinitis; ectopic dermatitis;  
 XX KW Der f I; Der p I; Der p II; Der f II.  
 XX OS Dermatophagoides sp.  
 XX PN US5968526-A.  
 XX PD 19-OCT-1999.  
 XX PF 07-JUN-1995; 95US-0478572.  
 XX PR 19-MAY-1995; 95US-0445307.  
 XX PR 14-APR-1994; 94US-0227772.  
 XX PR 12-APR-1995; 95WO-US04481.  
 XX PA (IMMU-) IMMULOGIC PHARM CORP.  
 XX PI Garman RD, Greenstein JL, Rogers BL, Franzen HM, Shaked Z, Chen X;  
 XX PI Evans S, Kuo M;  
 XX WPI: 1999-590385/50.  
 XX DT Screening individuals for allergic reactions to T cell epitopes of  
 XX PT major allergens from house dust mites -  
 XX PS Disclosure: Column 101-102; 158pp; English.  
 XX CC This invention describes a novel method (I) for detecting whether an  
 CC individual is sensitive to Dermatophagoides (house dust mites). The  
 CC method involves detecting sensitivity to house dust mites in patients,  
 CC comprising combining a blood sample from the individual with 1 or more  
 CC isolated T cell epitopes of the protein allergens I and II ((DP I) and  
 CC ((DP II)) from Dermatophagoides (house dust mites). 32 T cell epitopes  
 CC with varying, defined amino acids sequences (given in the specification)  
 CC may be used in (I). The sample and allergens are combined under  
 CC conditions appropriate for the binding of blood components with the  
 CC polypeptides. The extent of binding is then indicative of the  
 CC sensitivity of the patient to house dust mites. (I) may be used to screen  
 CC individuals for sensitivity to Dermatophagoides (house dust mites). The  
 CC house dust mite is a major cause of a variety of allergic disorders such  
 CC as asthma, rhinitis and ectopic dermatitis. AAY50360-Y50542 and  
 CC AAY50546-Y50555 represent house dust mite allergen peptide fragments  
 CC derived from Der p I, Der f II, Der f I and Der f II.  
 XX Sequence 25 AA;  
 Query Match 50.8%; Score 92; DB 20; Length 25;  
 Best Local Similarity 93.8%; Pred. No. 3.1e-06;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 17 GCHGSEPCNIHRGKPF 32  
 DB 10 GCHGSEPCIIHRGKPF 25  
 ||||| |||||

RESULT 8  
 AAU18996  
 ID AAU18996 standard; Peptide; 25 AA.  
 XX AC AAU18996;  
 XX DT 04-DEC-2001 (first entry)  
 XX DE T-cell epitope containing peptide DP11-2.  
 XX

KW House dust mite; allergenic peptide; Der p I; Der p II; Der f I;  
 KW Der f II; antiallergenic; immunostimulant; house dust mite allergy;  
 KW T-cell epitope.

XX Dermatophagoides pteronyssinus.

XX US6268491-B1.

XX 31-JUL-2001.

XX 07-JUN-1995; 95US-0484296.

XX 19-MAY-1995; 95US-0445307.

XX 16-OCT-1991; 91US-0777859.

XX 08-MAY-1992; 92US-0881396.

XX 14-APR-1993; 93WO-US03471.

XX 14-APR-1994; 94US-0227772.

XX (IMMU-) IMMULOGIC PHARM CORP.

XX Garman RD, Greenstein JL, Kuo M, Rogers BL, Franzen HM, Chen X;  
 PI Evans S, Shaked Z;

XX WPI; 2001-549074/61.

XX Peptides comprising T cell groups of the major allergens from  
 PT Dermatophagoides (house dust mites), useful for treating house dust  
 PT mite allergy in humans, and for diagnosing sensitivity to house dust  
 PT mite protein allergens -

XX Claim 5; Figure 3; 158pp; English.

XX The invention relates to an isolated peptide of the major protein  
 CC allergens of the genus Dermatophagoides, which comprises at least one T  
 CC cell group of a protein allergen from Der p (DP) I, DP II, Der f (DF) I  
 CC or DF II. The isolated peptide comprises at least two regions,  
 CC each region comprising at least one T cell group of a protein allergen  
 CC of the genus Dermatophagoides. The regions are derived from the same or  
 CC different protein allergens of the genus Dermatophagoides. The peptides  
 CC are useful for treating house dust mite allergy in humans. The peptides  
 CC are also useful for detecting or diagnosing sensitivity to house dust  
 CC mite protein allergens. The present peptides have similar or enhanced  
 CC therapeutic properties as the naturally-occurring allergen, but have  
 CC reduced side effects, and increased solubility and stability. The  
 CC present sequence represents an allergenic T-cell epitope containing  
 CC peptide derived from the Dermatophagoides allergenic proteins.

XX Sequence 25 AA;

Query Match 50.8%; Score 92; DB 22; Length 25;  
 Best Local Similarity 93.8%; Pred. No. 3.1e-06;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 17 GCHGSEPCNIHRGKPF 32

Db 10 GCHGSEPCNIHRGKPF 25

RESULT 9

AAR36478

ID AAR36478 standard; peptide; 25 AA.

XX AAR36478;

XX 12-AUG-1993 (first entry)

XX DFII-2(11-35), a Dermatophagoides protein allergen.

XX T cell epitope; house dust mite; allergy; soluble; Def pII.

XX Synthetic.

XX WO9308279-A.

PN

XX 29-APR-1993.

XX 15-OCT-1992; 92WO-US08637.

XX 16-OCT-1991; 91US-0777859.

XX 08-MAY-1992; 92US-0881396.

XX (IMMU-) IMMULOGIC PHARM CORP.

XX Garman RD, Greenstein JL, Kuo MC, Rogers BL;

XX WPI; 1993-152472/18.

XX Isolated peptide(s) of Dermatophagoides protein allergens - for  
 PT diagnosis and treatment of sensitivity to house dust mite

XX Claim 44; Fig 4; 176pp; English.

XX The peptide is one of a series of overlapping peptides synthesised by  
 CC standard techniques to cover the whole Dermatophagoides farinae  
 CC Def pII sequence. The T cell epitopes of the protein  
 CC were mapped by detection of the peptide's ability to stimulate T  
 CC cell activity. The peptides may be used for diagnosis and treatment  
 CC of sensitivity to house dust mite allergens. When administered to  
 CC house dust mite sensitive individuals, the peptides are capable of  
 CC modifying the allergic response to the allergens. The peptides may  
 CC be modified for e.g. increasing solubility, enhancing therapeutic or  
 CC preventive efficacy or stability.  
 CC See also AAR34686-700 and AAR36398-490.

XX Sequence 25 AA;

Query Match 49.2%; Score 89; DB 14; Length 25;  
 Best Local Similarity 87.5%; Pred. No. 8.4e-06;  
 Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 17 GCHGSEPCNIHRGKPF 32

Db 10 GCHGSEPCNIHRGKPF 25

RESULT 10

AAW72000

ID AAW72000 standard; peptide; 25 AA.

XX AAW72000;

XX 16-DEC-1998 (first entry)

XX Dermatophagoides Der f II protein peptide DFII-2.

XX genus Dermatophagoides; major protein allergen; T cell epitope;  
 KW Der p I; Der p II; Der f I; Der f II; house dust mite allergy.

XX Dermatophagoides sp.

XX US5820862-A.

XX 13-OCT-1998.

XX 07-JUN-1995; 95US-0482142.

XX 19-MAY-1995; 95US-0445307.

XX 14-APR-1994; 94US-0227772.

XX 07-JUN-1995; 95US-0482142.

XX (IMMU-) IMMULOGIC PHARM CORP.

XX Chen X, Evans S, Franzen HM, Garman RD, Greenstein JL;

XX Kuo M, Rogers BL, Shaked Z;

XX WPI; 1998-567590/48.

XX Dermatophagoides allergen peptides - useful for treating house dust  
PT mite allergy  
PS Disclosure: Column 129-130; 155pp; English.  
XX  
XX The present invention describes peptides for treating sensitivity to  
CC house dust mite allergens from the genus Dermatophagoides. Peptides  
CC within the scope of the invention comprise at least one T cell epitope,  
CC or preferably at least two T cell epitopes of a protein allergen  
CC selected from the allergens Der p I, Der p II, Der f I, or Der f II.  
CC The invention also describes modified peptides having similar or  
CC enhanced therapeutic properties as the corresponding, naturally  
CC occurring allergen, but having reduced side effects. AAW71912 to  
CC AAW72000, and AAW72257 to AAW72330 represent peptides from the present  
CC invention.  
XX  
XX Sequence 25 AA;  
SQ  
Query Match 49.2%; Score 89; DB 19; Length 25;  
Best Local Similarity 87.5%; Pred. No. 8.4e-06;  
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 17 GCHGSEPCNIHRGKPF 32  
DB 10 GCHGSDPCIHRGKPF 25  
|||||:|| |||||  
RESULT 11  
AAV50455  
ID AAV50455 standard; Peptide; 25 AA.  
XX  
AC AAV50455;  
XX  
DT 25-JAN-2000 (first entry)  
XX  
DE Dermatophagoides sp major protein allergen DF II-2.  
XX  
KW Allergen; house dust mite; detection; sensitivity; T cell epitope;  
KW screening; allergic disorder; asthma; rhinitis; ectopic dermatitis;  
KW Der f I; Der p I; Der p II; Der f II.  
XX  
OS Dermatophagoides sp.  
XX  
PN US5968526-A.  
XX  
PD 19-OCT-1999.  
XX  
PF 07-JUN-1995; 95US-0478572.  
XX  
PR 19-MAY-1995; 95US-0445307.  
PR 14-APR-1994; 94US-0227772.  
PR 12-APR-1995; 93WO-US04481.  
XX  
PA (IMMU-) IMMULOGIC PHARM CORP.  
XX  
PI Garman RD, Greenstein JL, Rogers BL, Franzen HM, Shaked Z, Chen X;  
PI Evans S, Kuo M;  
XX  
DR WPI; 1999-590385/50.  
XX  
PT Screening individuals for allergic reactions to T cell epitopes of  
PT major allergens from house dust mites -  
XX  
XX Claim 3m'; Column 131-132; 158pp; English.  
XX  
XX This invention describes a novel method (I) for detecting whether an  
CC individual is sensitive to Dermatophagoides (house dust mites). The  
CC method involves detecting sensitivity to house dust mites in patients,  
CC comprising combining a blood sample from the individual with 1 or more  
CC isolated T cell epitopes of the protein allergens I and II ((DP I) and  
CC (DP II)) from Dermatophagoides (house dust mites). 32 T cell epitopes  
CC with varying, defined amino acids sequences (given in the specification)

CC may be used in (I). The sample and allergens are combined under  
CC conditions appropriate for the binding of blood components with the  
CC polypeptides. The extent of binding is then indicative of the  
CC sensitivity of the patient to house dust mites. (I) may be used to screen  
CC individuals for sensitivity to Dermatophagoides (house dust mites). The  
CC house dust mite is a major cause of a variety of allergic disorders such  
CC as asthma, rhinitis and ectopic dermatitis. AAY50360-Y50542 and  
CC AAY50546-Y50555 represent house dust mite allergen peptide fragments  
CC derived from Der p I, Der f II, Der f I and Der f II.  
XX  
SQ Sequence 25 AA;  
Query Match 49.2%; Score 89; DB 20; Length 25;  
Best Local Similarity 87.5%; Pred. No. 8.4e-06;  
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 17 GCHGSEPCNIHRGKPF 32  
DB 10 GCHGSDPCIHRGKPF 25  
|||||:|| |||||  
RESULT 12  
AAU19058  
ID AAU19058 standard; Peptide; 25 AA.  
XX  
AC AAU19058;  
XX  
DT 04-DEC-2001 (first entry)  
XX  
DE T-cell epitope containing peptide DFII-2.  
XX  
KW House dust mite; allergenic peptide; Der p I; Der p II; Der f I;  
KW Der f II; anti-allergenic; immunostimulant; house dust mite allergy;  
KW T-cell epitope.  
XX  
OS Dermatophagoides farinae.  
XX  
PN US6268491-B1.  
XX  
PD 31-JUL-2001.  
XX  
PF 07-JUN-1995; 95US-0484296.  
XX  
PR 19-MAY-1995; 95US-0445307.  
PR 16-OCT-1991; 91US-0777859.  
PR 08-MAY-1992; 92US-0881396.  
PR 14-APR-1993; 93WO-US03471.  
PR 14-APR-1994; 94US-0227772.  
XX  
PA (IMMU-) IMMULOGIC PHARM CORP.  
XX  
PI Garman RD, Greenstein JL, Kuo M, Rogers BL, Franzen HM, Chen X;  
PI Evans S, Shaked Z;  
XX  
DR WPI; 2001-549074/61.  
XX  
PT Peptides comprising T cell groups of the major allergens from  
PT Dermatophagoides (house dust mites), useful for treating house dust  
PT mite allergy in humans, and for diagnosing sensitivity to house dust  
PT mite protein allergens -  
XX  
XX Claim 2; Column 131; 158pp; English.  
XX  
XX The invention relates to an isolated peptide of the major protein  
CC allergens of the genus Dermatophagoides, which comprises at least one T  
CC cell group of a protein allergen from Der p (DP) I, DP II, Der f (DF) I  
CC or DF II. The isolated peptide comprises at least two regions.  
CC each region comprising at least one T cell group of a protein allergen  
CC of the genus Dermatophagoides. The regions are derived from the same or  
CC different protein allergens of the genus Dermatophagoides. The peptides  
CC are useful for treating house dust mite allergy in humans. The peptides  
CC are also useful for detecting or diagnosing sensitivity to house dust  
CC mite protein allergens. The present peptides have similar or enhanced



CC therapeutic properties as the naturally-occurring allergen, but have  
 CC reduced side effects, and increased solubility and stability. The  
 CC present sequence represents an allergenic T-cell epitope containing  
 CC peptide derived from the Dermatophagoides allergenic proteins.  
 XX  
 SO Sequence 25 AA;  
 Query Match 49.2%; Score 89; DB 22; Length 25;  
 Best Local Similarity 87.5%; Pred. No. 8.4e-06;  
 Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Qy 17 GCHGSEPCNIHRGKPF 32  
 | | | | : | | | | | | | |  
 Db 10 GCHGSDPCIIHRGKPF 25  
 RESULT 13  
 ID AAB49092 standard; Protein; 25 AA.  
 XX AC AAB49092;  
 XX 27-MAR-2001 (first entry)  
 XX  
 XX Amyloid beta/tetanus toxoid immunogenic fusion protein, SEQ ID NO:28.  
 XX  
 KW Amyloid disease; amyloid fibril deposition; amyloid plaque;  
 KW immunogenic; antibody; vaccine; Alzheimer's disease;  
 KW type 2 diabetes; reactive system amyloidosis;  
 KW systemic senile amyloidosis; familial amyloid cardiomyopathy;  
 KW transmissible spongiform encephalopathy; Creutzfeld-Jakob disease; Kuru;  
 KW haemodialysis-associated beta-2-microglobulin deposition;  
 KW amyloid beta peptide; universal T-cell epitope; neuroprotective.  
 XX  
 OS Chimeric - Homo sapiens.  
 OS Chimeric - Clostridium tetani.  
 XX  
 PN WO200072876-A2.  
 XX 07-DEC-2000.  
 XX  
 XX 01-JUN-2000; 2000WO-US15239.  
 XX  
 XX 01-JUN-1999; 99US-0137010.  
 XX  
 XX (NEUR-) NEURALAB LTD.  
 XX  
 XX Schenk DB;  
 XX  
 XX WPI: 2001-070921/08.  
 XX  
 XX Pharmaceutical composition comprising immunogen against amyloid  
 PT component such as fibril peptide or protein, or antibody against  
 PT amyloid component useful for treating amyloid diseases or amyloidoses -  
 XX  
 XX Disclosure; Page 46; 140pp; English.  
 XX  
 XX The invention relates to a novel pharmaceutical composition for  
 CC preventing or treating a disease characterised by amyloid fibril  
 CC deposits (amyloid plaques) in a patient. The pharmaceutical composition  
 CC comprises an agent that will induce an immune response against an amyloid  
 CC component, or an antibody or antibody fragment that binds to an amyloid  
 CC component. The invention also relates to a method for determining  
 CC the prognosis of a patient undergoing treatment for an amyloid disorder  
 CC which involves measuring a patient serum amount of immunoreactivity  
 CC against a selected amyloid component. A patient serum immunoreactivity  
 CC of at least four times a base line serum immunoreactivity control level  
 CC indicates a prognosis of improved status with respect to the disorder.  
 CC The pharmaceutical compositions of the invention are useful for treating  
 CC a wide variety of disorders characterised by amyloid fibril deposition in  
 CC a patient. Such disorders include Alzheimer's disease characterised by  
 CC amyloid beta peptide fibril deposits; type 2 diabetes characterised by  
 CC islet amyloid protein peptide (IAPP, amylin) fibrils; reactive systemic

CC amyloidosis associated with systemic inflammatory diseases (e.g.,  
 CC rheumatoid arthritis, osteomyelitis, tuberculosis) characterised by AA  
 CC fibrils derived from serum amyloid A protein (ApoSAA)); systemic senile  
 CC amyloidosis and familial amyloid cardiomyopathy characterised by ATTR  
 CC fibrils derived from transthyretin (TTR); transmissible spongiform  
 CC encephalopathies (e.g. Creutzfeld-Jakob disease, Kuru) characterised by  
 CC prion protein deposits; and beta-2-microglobulin deposits which form as  
 CC a result of long term haemodialysis treatment. The present sequence  
 CC represents an immunogenic fusion protein comprising an amyloid beta  
 CC peptide fused to a universal T-cell epitope which may be used in a  
 CC composition to treat or prevent Alzheimer's disease.  
 XX  
 XX Sequence 25 AA;  
 Query Match 48.1%; Score 87; DB 22; Length 25;  
 Best Local Similarity 94.4%; Pred. No. 1.6e-05;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 QYIKANSKFIGITELGKC 18  
 | | | | | | | | | | | | | |  
 Db 8 QYIKANSKFIGITELGKC 25  
 RESULT 14  
 AAY82638  
 ID AAY82638 standard; peptide; 15 AA.  
 XX AC AAY82638;  
 XX 07-AUG-2000 (first entry)  
 XX  
 XX Der pII peptide SEQ ID NO:14.  
 XX  
 KW T cell epitope; B cell epitope; allergy; allergen; antigenic;  
 KW anti-allergic; antiasthmatic; anti-inflammatory; dermatological;  
 KW immunosuppressive; vaccine; rhinitis; sinusitis; bronchial asthma;  
 KW atopic dermatitis; acute urticaria; chronic urticaria;  
 KW gastro-intestinal syndrome; food allergen; oro-pharyngeal syndrome;  
 KW anaphylactic reaction; drug hypersensitivity; allergic reaction.  
 XX  
 OS Dermatophagoides pteronyssinus.  
 OS Synthetic.  
 XX  
 XX WO200006694-A2.  
 XX  
 XX 10-FEB-2000.  
 XX  
 XX 20-JUL-1999; 99WO-BE00092.  
 XX  
 XX 30-JUL-1998; 98EP-0870167.  
 XX  
 XX (UNIO ) UCB SA.  
 XX  
 XX Saint-Remy J, Jacquemin M;  
 XX  
 XX WPI: 2000-422470/36.  
 XX  
 XX New compound for prevention and treatment of allergies comprises at  
 PT least one allergen antigenic determinant recognized by a B cell and at  
 PT least one antigenic determinant which does not trigger T cell  
 PT activation -  
 XX  
 XX Example 2; Page 27; 50pp; English.  
 XX  
 CC The present invention describes a compound (I) for the prevention and/or  
 CC treatment of allergy. The compound comprises at least one allergen  
 CC antigenic determinant (I) recognised by a B cell or an antibody secreted  
 CC by a B cell of a non-atopic individual and at least one antigenic  
 CC determinant (II) different from the allergen that triggers T cell  
 CC activation. (I) has anti-allergic, antiasthmatic, anti-inflammatory,  
 CC dermatological and immunosuppressive activities, and can be used in a  
 CC vaccine. (II) may be used in a pharmaceutical or cosmetic medicament to  
 CC treat and/or prevent allergies or a disease of allergic origin,

CC especially hypersensitivities. These include rhinitis, sinusitis,  
CC bronchial asthma, atopic dermatitis, some forms of acute and chronic  
CC urticaria, gastro-intestinal syndromes associated with the ingestion of  
CC food allergens, oro-pharyngeal syndrome, anaphylactic reactions  
CC associated with drug hypersensitivities and/or a mixture of these. The  
CC use of (I) in the treatment of allergic conditions avoids the need for  
CC drug treatment, which often causes undesirable side-effects. Also, prior  
CC art drug therapies alleviate symptoms, but do not influence their  
CC causes, however (I) actually combats the cause of an allergic reaction.  
CC The present sequence represents a peptide, which is used in an  
CC example from the present invention.

XX Sequence 15 AA;

Query Match 47.5%; Score 86; DB 21; Length 15;  
Best Local Similarity 93.3%; Pred. No. 1.3e-05;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 18 CHGSEPCNIHRGKPF 32  
||||| |||||  
Db 1 CHGSEPCIIHRGKPF 15

RESULT 15  
AAR83561  
ID AAR83561 standard; peptide; 29 AA.

AC AAR83561;

XX 13-JUN-1996 (first entry)

DE IgE CH4 region contg. peptide immunogen for treating allergies.

XX IgE; CH4; immunoglobulin; epsilon; immunogen; helper T cell; epitope;  
KW vaccine; allergy; antibody; constant heavy chain.

XX Synthetic.

XX WO9526365-A1.

XX 05-OCT-1995.

XX 24-MAR-1995; 95WO-US03741.

XX 25-OCT-1994; 94US-0328912.

XX 28-MAR-1994; 94US-0218461.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY;

XX WPI; 1995-351297/45.

XX Synthetic peptide-based immunogen contg. IgE CH4 peptide and helper  
XX T cell epitope - useful for eliciting antibody prodn. for allergy  
XX treatment

XX Claim 5; Page 68-69; 87pp; English.

XX AAR82592-R82600 and AAR83560-R83581 are peptide immunogens that are  
XX useful in vaccines for treating allergic reactions. In the immunogens,  
XX an IgE CH4 peptide is attached C-terminally to a series of amino acids  
XX including a helper T cell epitope. The immunogen may also opt. contain  
XX a fatty acid or fatty acid derivative, an invasin domain or alpha-NH2.  
XX The immunogen produces high titres of antibodies to the effector site  
XX in human IgE heavy chain (the CH4 domain peptide) which inhibit mast  
XX cell activation and reduce allergen-induced IgE prodn. The immunogens  
XX may be used in either a radially branching multimeric form or a  
XX linearly arranged monomeric form.

XX Sequence 29 AA;

Query Match 47.5%; Score 86; DB 16; Length 29;

Best Local Similarity 100.0%; Pred. No. 2.7e-05;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OYIKANSKFIGITELGG 17  
||||| |||||  
Db 3 OYIKANSKFIGITELGG 19

Search completed: December 4, 2002, 13:16:51  
Job time : 36 secs

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 13:17:47 ; Search time 11 Seconds  
(without alignments)  
47.250 Million cell updates/sec

Title: US-09-362-731A-5

Perfect score: 181

Sequence: 1 QYIKANSFIGITELGGCHGSEPCNIHRGKPF 32

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 40818

Minimum DB seq length: 0

Maximum DB seq length: 32

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*

```
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80	44.2	28	10	US-09-848-834A-11
2	76	42.0	31	10	US-09-943-548-2
3	74	40.9	15	10	US-09-862-849-2
4	74	40.9	16	10	US-09-848-834A-2
5	74	40.9	29	10	US-09-732-754-1
6	74	40.9	31	10	US-09-848-834A-15
7	70	38.7	29	10	US-09-983-019-8
8	70	38.7	29	10	US-09-983-019-9
9	70	38.7	31	10	US-09-983-019-3
10	70	38.7	31	10	US-09-983-019-5
11	70	38.7	31	10	US-09-983-019-6
12	58.5	32.3	14	9	US-10-044-034-21
13	40	22.1	17	10	US-09-864-761-47907
14	39	21.5	19	10	US-09-899-376-3
15	37	20.4	19	10	US-09-846-729A-24
16	37	20.4	19	10	US-09-846-729A-28
17	37	20.4	30	10	US-09-864-761-44068
18	36	19.9	28	9	US-09-749-637A-310
19	34	18.8	23	10	US-09-864-761-35627

Query Match 44.2%; Score 80; DB 10; Length 28;  
Best Local Similarity 100.0%; Pred. No. 1.5e-05;

20	34	18.8	30	10	US-09-894-882-481	Sequence 481, Appl
21	33.5	18.5	23	10	US-09-864-761-36005	Sequence 36005, A
22	33.5	18.5	24	8	US-08-424-550B-587	Sequence 587, App
23	33.5	18.5	31	8	US-08-424-550B-350	Sequence 350, App
24	33	18.2	8	10	US-09-897-107-23	Sequence 23, Appl
25	33	18.2	12	10	US-09-784-887B-6	Sequence 6, Appl
26	33	18.2	29	10	US-09-739-907-143	Sequence 143, App
27	32.5	18.0	28	10	US-09-864-761-47156	Sequence 47156, A
28	32	17.7	12	10	US-09-899-376-1	Sequence 1, Appl
29	32	17.7	12	10	US-09-784-887B-5	Sequence 5, Appl
30	32	17.7	15	9	US-09-851-026-24	Sequence 24, Appl
31	32	17.7	19	10	US-09-899-376-4	Sequence 4, Appl
32	32	17.7	20	10	US-09-864-761-34747	Sequence 34747, A
33	32	17.7	21	10	US-09-864-761-46555	Sequence 46555, A
34	32	17.7	22	9	US-09-310-113-4	Sequence 4, Appl
35	32	17.7	25	10	US-09-864-761-33647	Sequence 33647, A
36	32	17.7	26	10	US-09-864-761-35140	Sequence 35140, A
37	32	17.7	28	10	US-09-864-761-46263	Sequence 46263, A
38	31.5	17.4	29	9	US-09-749-637A-253	Sequence 253, App
39	31.5	17.4	29	10	US-09-864-761-46090	Sequence 46090, A
40	31.5	17.4	30	10	US-09-864-761-36210	Sequence 36210, A
41	31.5	17.4	31	10	US-09-864-761-37948	Sequence 37948, A
42	31	17.1	10	10	US-09-822-827-962	Sequence 962, App
43	31	17.1	21	9	US-09-310-113-17	Sequence 17, Appl
44	31	17.1	22	9	US-09-310-113-3	Sequence 3, Appl
45	31	17.1	22	9	US-09-310-113-18	Sequence 18, Appl

#### ALIGNMENTS

RESULT 1  
US-09-848-834A-11  
; Sequence 11, Application US/09848834A  
; Patent No. US20020076416A1  
; GENERAL INFORMATION:  
; APPLICANT: Apton Corporation  
; TITLE OF INVENTION: Chimeric Peptide Immunogens  
; FILE REFERENCE: 1102865-0047  
; CURRENT APPLICATION NUMBER: US/09/848,834A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/202,328  
; PRIOR FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 11  
; LENGTH: 28  
; TYPE: PPT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 830-844 of  
; OTHER INFORMATION: Tetanus toxoid precursor (Tentoxylisin) linked by a spacer to  
; OTHER INFORMATION: ino acid sequence 2-10 of the GnRH hormone  
; NAME/KEY: MOD\_RES  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Amidated-glutamine  
; NAME/KEY: MOD\_RES  
; LOCATION: (28)..(28)  
; OTHER INFORMATION: Amidated-glycine or glycineamide  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(15)  
; OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus Toxoid Precursor  
; OTHER INFORMATION: (Tentoxylisin)  
; NAME/KEY: PEPTIDE  
; LOCATION: (16)..(19)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (20)..(28)  
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone  
; US-09-848-834A-11



## RESULT 6

US-09-848-834A-15  
 ; Sequence 15, Application US/09848834A  
 ; Patent No. US20020076416A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Aphton Corporation  
 ; TITLE OF INVENTION: Chimeric Peptide Immunogens  
 ; FILE REFERENCE: 1102865-0047  
 ; CURRENT APPLICATION NUMBER: US/09/848,834A  
 ; CURRENT FILING DATE: 2001-05-04  
 ; PRIOR APPLICATION NUMBER: 60/202,328  
 ; PRIOR FILING DATE: 2000-05-05  
 ; NUMBER OF SEQ ID NOS: 20  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 15  
 ; LENGTH: 31  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the G  
 ; OTHER INFORMATION: RH hormone linked by a spacer to amino sequence 830-844 of the Te  
 ; OTHER INFORMATION: tanus toxoid precursor (Tentoxylisin)  
 ; NAME/KEY: MOD\_RES  
 ; LOCATION: (1)..(1)  
 ; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline  
 ; NAME/KEY: PEPTIDE  
 ; LOCATION: (1)..(10)  
 ; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone  
 ; NAME/KEY: PEPTIDE  
 ; LOCATION: (11)..(16)  
 ; OTHER INFORMATION: Spacer peptide  
 ; NAME/KEY: PEPTIDE  
 ; LOCATION: (17)..(31)  
 ; OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus toxoid precursor  
 ; OTHER INFORMATION: (Tentoxylisin)  
 US-09-848-834A-15

Query Match 40.9%; Score 74; DB 10; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 0.00013;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYKANSKFIGITEL 15  
 Db 17 QYKANSKFIGITEL 31

## RESULT 7

US-09-983-019-8  
 ; Sequence 8, Application US/09983019  
 ; Patent No. US20020146820A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Diamond, Don J.  
 ; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPTIOTES OF HUMAN CYTOMEGALOVIRUS PP15  
 ; FILE REFERENCE: 1954-347US  
 ; CURRENT APPLICATION NUMBER: US/09/983,019  
 ; CURRENT FILING DATE: 2001-10-22  
 ; PRIOR APPLICATION NUMBER: US 60/241,944  
 ; PRIOR FILING DATE: 2000-10-20  
 ; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 8  
 ; LENGTH: 29  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)..(1)  
 ; OTHER INFORMATION: Human cytomegalovirus vaccine peptide  
 ; OTHER INFORMATION: Human cytomegalovirus vaccine peptide  
 US-09-983-019-8

Query Match 38.7%; Score 70; DB 10; Length 29;  
 Best Local Similarity 100.0%; Pred. No. 0.00046;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYKANSKFIGITE 14  
 Db 16 QYKANSKFIGITE 29

## RESULT 8

US-09-983-019-9  
 ; Sequence 9, Application US/09983019  
 ; Patent No. US20020146820A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Diamond, Don J.  
 ; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPTIOTES OF HUMAN CYTOMEGALOVIRUS P  
 ; FILE REFERENCE: 1954-347US  
 ; CURRENT APPLICATION NUMBER: US/09/983,019  
 ; CURRENT FILING DATE: 2001-10-22  
 ; PRIOR APPLICATION NUMBER: US 60/241,944  
 ; PRIOR FILING DATE: 2000-10-20  
 ; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 9  
 ; LENGTH: 29  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)..(1)  
 ; OTHER INFORMATION: Human cytomegalovirus vaccine peptide  
 ; OTHER INFORMATION: Human cytomegalovirus vaccine peptide  
 US-09-983-019-9

Query Match 38.7%; Score 70; DB 10; Length 29;  
 Best Local Similarity 100.0%; Pred. No. 0.00046;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYKANSKFIGITE 14  
 Db 1 QYKANSKFIGITE 14

## RESULT 9

US-09-983-019-3  
 ; Sequence 3, Application US/09983019  
 ; Patent No. US20020146820A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Diamond, Don J.  
 ; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPTIOTES OF HUMAN CYTOMEGALOVIRUS P  
 ; FILE REFERENCE: 1954-347US  
 ; CURRENT APPLICATION NUMBER: US/09/983,019  
 ; CURRENT FILING DATE: 2001-10-22  
 ; PRIOR APPLICATION NUMBER: US 60/241,944  
 ; PRIOR FILING DATE: 2000-10-20  
 ; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 3  
 ; LENGTH: 31  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)..(1)  
 ; OTHER INFORMATION: Human cytomegalovirus vaccine peptide  
 ; NAME/KEY: LIPID  
 ; LOCATION: (1)..(1)  
 ; OTHER INFORMATION: di-palmitic acid  
 US-09-983-019-3

Query Match 38.7%; Score 70; DB 10; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 0.00049;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYKANSKFIGITE 14  
 Db 4 QYKANSKFIGITE 17

RESULT 10  
US-09-983-019-5  
; Sequence 5, Application US/09983019  
; Patent No. US20020146820A1  
; GENERAL INFORMATION:  
; APPLICANT: Diamond, Don J.  
; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS PP19  
; FILE REFERENCE: 1954-347US  
; CURRENT APPLICATION NUMBER: US/09/983,019  
; CURRENT FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: US 60/241,944  
; PRIOR FILING DATE: 2000-10-20  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 5  
; LENGTH: 31  
; TYPE: PPT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Human cytomegalovirus vaccine peptide  
; NAME/KEY: LIPID  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: palmitic acid  
US-09-983-019-5

Query Match 38.7%; Score 70; DB 10; Length 31;  
Best Local Similarity 100.0%; Pred. No. 0.00049;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 QYIKANSKFIGITE 14  
Db 4 QYIKANSKFIGITE 17

RESULT 11  
US-09-983-019-6  
; Sequence 6, Application US/09983019  
; Patent No. US20020146820A1  
; GENERAL INFORMATION:  
; APPLICANT: Diamond, Don J.  
; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS PP19  
; FILE REFERENCE: 1954-347US  
; CURRENT APPLICATION NUMBER: US/09/983,019  
; CURRENT FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: US 60/241,944  
; PRIOR FILING DATE: 2000-10-20  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 6  
; LENGTH: 31  
; TYPE: PPT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Human cytomegalovirus vaccine peptide  
; NAME/KEY: LIPID  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: palmitic acid  
US-09-983-019-6

Query Match 38.7%; Score 70; DB 10; Length 31;  
Best Local Similarity 100.0%; Pred. No. 0.00049;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 QYIKANSKFIGITE 14  
Db 18 QYIKANSKFIGITE 31

RESULT 12  
US-10-044-034-21  
; Sequence 21, Application US/10044034  
; Patent No. US20020169264A1  
; GENERAL INFORMATION:  
; APPLICANT: JACKSON, DAVID C.  
; APPLICANT: O'BRIEN-SIMPSON, NEIL M.  
; APPLICANT: BROWN, LORENA E.  
; APPLICANT: EDE, NICHOLAS J.  
; APPLICANT: BRANDT, EVELYN R.  
; APPLICANT: GOOD, MICHAEL F.  
; TITLE OF INVENTION: POLYMERS INCORPORATING PEPTIDES  
; FILE REFERENCE: FBRC:006  
; CURRENT APPLICATION NUMBER: US/10/044,034  
; CURRENT FILING DATE: 2002-01-11  
; PRIOR APPLICATION NUMBER: P05071  
; PRIOR FILING DATE: 1997-02-11  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 21  
; LENGTH: 14  
; TYPE: PPT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Peptides  
US-10-044-034-21

Query Match 32.3%; Score 58.5; DB 9; Length 14;  
Best Local Similarity 93.3%; Pred. No. 0.01;  
Matches 14; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
  
Qy 1 QYIKANSKFIGITEL 15  
Db 1 QYIKANS-FIGITEL 14

RESULT 13  
US-09-864-761-47907  
; Sequence 47907, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aeonica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47907
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000365.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.99
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
; OTHER INFORMATION: EST_HUMAN HIT: A0126316.1, EVALUATE 1.00e-04
US-09-864-761-47907
```

```
Query Match 22.1%; Score 40; DB 10; Length 17;
Best Local Similarity 52.4%; Pred. No. 6.3;
Matches 11; Conservative 1; Mismatches 3; Indels 6; Gaps 3;
```

```
Qy 9 FIGITELGGCHGSEPC--NIH 27
   | | | | | | | | | |
Db 1 PPGLT---GCH-SRCPNFEH 17
```

```
RESULT 14
US-09-899-376-3
; Sequence 3, Application US/09899376
; Patent No. US20020102265A1
; GENERAL INFORMATION:
; APPLICANT: HONG, FRANK D.
; APPLICANT: CLAYMAN, GARY
; TITLE OF INVENTION: ISOLATION OF A CELL-SPECIFIC INTERNALIZING PEPTIDE THAT
; TITLE OF INVENTION: INFILTRATES TUMOR TISSUE FOR TARGETED DRUG DELIVERY
; FILE REFERENCE: UTSC:645US
; CURRENT APPLICATION NUMBER: US/09/899,376
; CURRENT FILING DATE: 2001-07-02
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-09-899-376-3
```

```
Query Match 21.5%; Score 39; DB 10; Length 19;
Best Local Similarity 53.3%; Pred. No. 9.9;
Matches 8; Conservative 2; Mismatches 3; Indels 2; Gaps 1;
```

```
Qy 16 GGCHGSEPCNIHRGK 30
   | | | | | | | | | |
Db 1 GG--GTSPLNIHQ 13
```

```
RESULT 15
US-09-846-729A-24
; Sequence 24, Application US/09846729A
```

```
; Patent No. US20020058322A1
; GENERAL INFORMATION:
; APPLICANT: Boone, Thomas
; APPLICANT: Li, Huimin
; APPLICANT: Mann, Michael
; TITLE OF INVENTION: FIBRINOLYTICALLY ACTIVE POLYPEPTIDE
; FILE REFERENCE: A-596
; CURRENT APPLICATION NUMBER: US/09/846,729A
; CURRENT FILING DATE: 2001-05-01
; PRIOR APPLICATION NUMBER: 09/411,329
; PRIOR FILING DATE: 1999-10-01
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Agkistrodon contortrix
; NAME/KEY: misc.feature
; FEATURE:
; OTHER INFORMATION: Native pro-fibrolase of Agkistrodon contortrix
US-09-846-729A-24
```

```
Query Match 20.4%; Score 37; DB 10; Length 19;
Best Local Similarity 37.5%; Pred. No. 19;
Matches 6; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
```

```
Qy 12 ITELGGCHGSEPCNIH 27
   | | | | | | | | | |
Db 3 IKRMSECHLPERCLRH 18
```

```
Search completed: December 4, 2002, 13:21:30
Job time : 12 secs
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: December 4, 2002, 13:16:12 : Search time 15 Seconds  
(without alignments)  
62.769 Million cell updates/sec

Title: US-09-362-731A-5  
Perfect score: 181  
Sequence: 1 QYIKANSKFIGITELGCGHSEPCNIHRGKPF 32

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 161394

Minimum DB seq length: 0  
Maximum DB seq length: 32

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/1aa/5A-COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B-COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A-COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B-COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PCITUS-COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92	50.8	25	US-08-482-142-42	Sequence 42, Appl
2	92	50.8	25	US-08-478-572-42	Sequence 42, Appl
3	92	50.8	25	US-08-484-296-42	Sequence 42, Appl
4	92	50.8	25	PCT-US95-04481-19	Sequence 19, Appl
5	89	49.2	25	US-08-482-142-104	Sequence 104, Appl
6	89	49.2	25	US-08-478-572-104	Sequence 104, Appl
7	89	49.2	25	US-08-484-296-104	Sequence 104, Appl
8	77	42.5	29	US-08-482-142-43	Sequence 43, Appl
9	77	42.5	29	US-08-482-142-119	Sequence 119, Appl
10	77	42.5	29	US-08-482-142-206	Sequence 206, Appl
11	77	42.5	29	US-08-478-572-43	Sequence 43, Appl
12	77	42.5	29	US-08-478-572-119	Sequence 119, Appl
13	77	42.5	29	US-08-478-572-206	Sequence 206, Appl
14	77	42.5	29	US-08-484-296-43	Sequence 43, Appl
15	77	42.5	29	US-08-484-296-119	Sequence 119, Appl
16	77	42.5	29	US-08-484-296-206	Sequence 206, Appl
17	77	42.5	29	PCT-US95-04481-20	Sequence 20, Appl
18	76	42.0	31	US-08-432-483A-2	Sequence 2, Appl
19	74	40.9	15	US-08-319-704-10	Sequence 10, Appl
20	74	40.9	15	US-08-661-052-6	Sequence 6, Appl
21	74	40.9	15	US-08-460-502-7	Sequence 7, Appl
22	74	40.9	15	US-08-046-373-2	Sequence 2, Appl
23	74	40.9	15	US-09-188-082-6	Sequence 6, Appl
24	74	40.9	15	US-09-364-088-6	Sequence 6, Appl
25	74	40.9	15	US-09-102-716-6	Sequence 6, Appl
26	74	40.9	15	US-09-148-711A-7	Sequence 7, Appl
27	74	40.9	15	US-08-716-249-4	Sequence 4, Appl

Sequence 69, Appl  
Sequence 55, Appl  
Sequence 4, Appl  
Sequence 4, Appl  
Sequence 23, Appl  
Sequence 7, Appl  
Sequence 13, Appl  
Sequence 13, Appl  
Sequence 105, Appl  
Sequence 105, Appl  
Sequence 105, Appl  
Sequence 63, Appl  
Sequence 5, Appl  
Sequence 18, Appl  
Sequence 9, Appl  
Sequence 15, Appl

Sequence 69, Appl  
Sequence 55, Appl  
Sequence 4, Appl  
Sequence 4, Appl  
Sequence 23, Appl  
Sequence 7, Appl  
Sequence 13, Appl  
Sequence 13, Appl  
Sequence 105, Appl  
Sequence 105, Appl  
Sequence 105, Appl  
Sequence 63, Appl  
Sequence 5, Appl  
Sequence 18, Appl  
Sequence 9, Appl  
Sequence 15, Appl

## ALIGNMENTS

RESULT 1  
US-08-482-142-42  
; Sequence 42, Application US/08482142  
; Patent No. 5820862  
; GENERAL INFORMATION:  
; APPLICANT: Garman, Richard  
; APPLICANT: Greenstein, Julia  
; APPLICANT: Kuo, Mei-chang  
; APPLICANT: Rogers, Bruce  
; APPLICANT: Franzen, Henry  
; APPLICANT: Chen, Xian  
; APPLICANT: Evans, Sean  
; APPLICANT: Shaked, Ze'ev  
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
; NUMBER OF SEQUENCES: 207  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
; STREET: 610 LINCOLN STREET  
; CITY: WALTHAM  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/482,142  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/445,307  
; FILING DATE: 07 June 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CRAIG, ANNE I.  
; REGISTRATION NUMBER: 32,976  
; REFERENCE/DOCKET NUMBER: 017.6US  
; TELEPHONE: (617) 466-6000  
; TELEFAX: (617) 466-6040  
; INFORMATION FOR SEQ ID NO: 42:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 25 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: N-terminal  
US-08-482-142-42



RESULT 3  
US-08-484-296-42  
; Sequence 42, Application US/08484296

```

;
; APPLICANT: Pharmaceutical Peptide Formulations
;
; TITLE OF INVENTION: 54
; NUMBER OF SEQUENCES: 54
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04481
; FILING DATE:
; CLASSIFICATION:
;

```

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/227,772  
;; FILING DATE: April 14, 1994  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Vanstone, Darlene A.  
;; REGISTRATION NUMBER: 35,279  
;; REFERENCE/DOCKET NUMBER: 017.5 PCT  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 466-6000  
;; TELEFAX: (617) 466-6040  
;; INFORMATION FOR SEQ ID NO: 19:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 25 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS:  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FRAGMENT TYPE: internal  
PCT-US95-04481-19

Query Match 50.8%; Score 92; DB 5; Length 25;  
Best Local Similarity 93.8%; Pred. No. 2.6e-06;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 17 GCHGSEPCNIHRGKPF 32  
|||||:|||||  
Db 10 GCHGSEPCNIHRGKPF 25

RESULT 5  
US-08-482-142-104  
;; Sequence 104, Application US/08482142  
;; Patent No. 5820862  
;; GENERAL INFORMATION:  
;; APPLICANT: Garman, Richard  
;; APPLICANT: Greenstein, Julia  
;; APPLICANT: Kuo, Mei-chang  
;; APPLICANT: Rogers, Bruce  
;; APPLICANT: Franzen, Henry  
;; APPLICANT: Chen, Xian  
;; APPLICANT: Evans, Sean  
;; APPLICANT: Shaked, Ze'ev  
;; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
;; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
;; NUMBER OF SEQUENCES: 207  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
;; STREET: 610 LINCOLN STREET  
;; CITY: WALTHAM  
;; STATE: MA  
;; COUNTRY: USA  
;; ZIP: 02154  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: ASCII TEXT  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/482,142  
;; FILING DATE: 07-JUN-1995  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/445,307  
;; FILING DATE: 07 June 1995  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: CRAIG, ANNE I.  
;; REGISTRATION NUMBER: 32,976  
;; REFERENCE/DOCKET NUMBER: 017.60S  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 466-6000  
;; TELEFAX: (617) 466-6040  
;; INFORMATION FOR SEQ ID NO: 104:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 25 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FRAGMENT TYPE: N-terminal  
US-08-478-572-104

Query Match 49.2%; Score 89; DB 2; Length 25;  
Best Local Similarity 87.5%; Pred. No. 6.7e-06;  
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 17 GCHGSEPCNIHRGKPF 32  
|||||:|||||

;; LENGTH: 25 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FRAGMENT TYPE: N-terminal  
US-08-482-142-104

Query Match 49.2%; Score 89; DB 2; Length 25;  
Best Local Similarity 87.5%; Pred. No. 6.7e-06;  
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 17 GCHGSEPCNIHRGKPF 32  
|||||:|||||  
Db 10 GCHGSEPCNIHRGKPF 25

RESULT 6  
US-08-478-572-104  
;; Sequence 104, Application US/08478572  
;; Patent No. 5968526  
;; GENERAL INFORMATION:  
;; APPLICANT: Garman, Richard  
;; APPLICANT: Greenstein, Julia  
;; APPLICANT: Kuo, Mei-chang  
;; APPLICANT: Rogers, Bruce  
;; APPLICANT: Franzen, Henry  
;; APPLICANT: Chen, Xian  
;; APPLICANT: Evans, Sean  
;; APPLICANT: Shaked, Ze'ev  
;; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
;; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
;; NUMBER OF SEQUENCES: 207  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
;; STREET: 610 LINCOLN STREET  
;; CITY: WALTHAM  
;; STATE: MA  
;; COUNTRY: USA  
;; ZIP: 02154  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: ASCII TEXT  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/478,572  
;; FILING DATE: 07-June-1995  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/445,307  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: CRAIG, ANNE I.  
;; REGISTRATION NUMBER: 32,976  
;; REFERENCE/DOCKET NUMBER: 017.60S  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 466-6000  
;; TELEFAX: (617) 466-6040  
;; INFORMATION FOR SEQ ID NO: 104:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 25 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FRAGMENT TYPE: N-terminal  
US-08-478-572-104

Query Match 49.2%; Score 89; DB 2; Length 25;  
Best Local Similarity 87.5%; Pred. No. 6.7e-06;  
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 10 GCHGSDPCIHRGKPF 25

## RESULT 7

US-08-484-296-104

; Sequence 104, Application US/08484296

; Patent No. 6268491

; GENERAL INFORMATION:

; APPLICANT: Garman, Richard

; APPLICANT: Greenstein, Julia

; APPLICANT: Kuo, Mei-chang

; APPLICANT: Rogers, Bruce

; APPLICANT: Franzen, Henry

; APPLICANT: Chen, Xian

; APPLICANT: Evans, Sean

; APPLICANT: Shaked, Ze'ev

; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS

; FROM DERMATOPHAGOIDES (HOUSE DUST MITE)

; NUMBER OF SEQUENCES: 207

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION

; STREET: 610 LINCOLN STREET

; CITY: WALTHAM

; STATE: MA

; COUNTRY: USA

; ZIP: 02154

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII TEXT

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/484,296

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/445,307

; FILING DATE: 07 June 1995

; ATTORNEY/AGENT INFORMATION:

; NAME: CRAIG, ANNE I.

; REGISTRATION NUMBER: 32,976

; REFERENCE/DOCKET NUMBER: 017.605

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 466-6000

; TELEFAX: (617) 466-6040

; INFORMATION FOR SEQ ID NO: 104:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 25 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; FRAGMENT TYPE: N-terminal

US-08-484-296-104

Query Match 49.2%; Score 89; DB 4; Length 25;

Best Local Similarity 87.5%; Pred. No. 6.7e-06;

Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 17 GCHGSDPCIHRGKPF 32

|||||:|||||

Db 10 GCHGSDPCIHRGKPF 25

## RESULT 8

US-08-482-142-43

; Sequence 43, Application US/08482142

; Patent No. 5820862

; GENERAL INFORMATION:

; APPLICANT: Garman, Richard

; APPLICANT: Greenstein, Julia

; APPLICANT: Kuo, Mei-chang

; APPLICANT: Rogers, Bruce

; APPLICANT: Franzen, Henry

; APPLICANT: Chen, Xian

; APPLICANT: Evans, Sean

; APPLICANT: Shaked, Ze'ev

; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS

; FROM DERMATOPHAGOIDES (HOUSE DUST MITE)

; NUMBER OF SEQUENCES: 207

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION

; STREET: 610 LINCOLN STREET

; CITY: WALTHAM

; STATE: MA

; COUNTRY: USA

; ZIP: 02154

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII TEXT

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/482,142

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/445,307

; FILING DATE: 07 June 1995

; ATTORNEY/AGENT INFORMATION:

; NAME: CRAIG, ANNE I.

; REGISTRATION NUMBER: 32,976

; REFERENCE/DOCKET NUMBER: 017.605

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 466-6000

; TELEFAX: (617) 466-6040

; INFORMATION FOR SEQ ID NO: 43:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 29 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; FRAGMENT TYPE: N-terminal

US-08-482-142-43

Query Match 42.5%; Score 77; DB 2; Length 29;

Best Local Similarity 92.9%; Pred. No. 0.00035;

Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 19 HGSEPCIIHRGKPF 32

|||||:|||||

Db 1 HGSEPCIIHRGKPF 14

## RESULT 9

US-08-482-142-119

; Sequence 119, Application US/08482142

; Patent No. 5820862

; GENERAL INFORMATION:

; APPLICANT: Garman, Richard

; APPLICANT: Greenstein, Julia

; APPLICANT: Kuo, Mei-chang

; APPLICANT: Rogers, Bruce

; APPLICANT: Franzen, Henry

; APPLICANT: Chen, Xian

; APPLICANT: Evans, Sean

; APPLICANT: Shaked, Ze'ev

; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS

; FROM DERMATOPHAGOIDES (HOUSE DUST MITE)

; NUMBER OF SEQUENCES: 207

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION

; STREET: 610 LINCOLN STREET

; CITY: WALTHAM

; STATE: MA

; COUNTRY: USA

; ZIP: 02154

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/482,142  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/445,307  
; FILING DATE: 07 June 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CRAIG, ANNE I.  
; REGISTRATION NUMBER: 32,976  
; REFERENCE/DOCKET NUMBER: 017.60US  
; TELEPHONE: (617) 466-6000  
; TELEFAX: (617) 466-6040  
; INFORMATION FOR SEQ ID NO: 119:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 29 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: N-terminal  
; US-08-482-142-119

Query Match 42.5%; Score 77; DB 2; Length 29;  
Best Local Similarity 92.9%; Pred. No. 0.00035;  
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 19 HGSEPCNIHRGKPF 32  
||||| |||||  
Db 1 HGSEPCNIHRGKPF 14

RESULT 10  
US-08-482-142-206  
; Sequence 206, Application US/08482142  
; Patent No. 5820862  
; GENERAL INFORMATION:  
; APPLICANT: Garman, Richard  
; APPLICANT: Greenstein, Julia  
; APPLICANT: Kuo, Mei-chang  
; APPLICANT: Rogers, Bruce  
; APPLICANT: Franzen, Henry  
; APPLICANT: Chen, Xian  
; APPLICANT: Evans, Sean  
; APPLICANT: Shaked, Ze'ev  
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
; NUMBER OF SEQUENCES: 207  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
; STREET: 610 LINCOLN STREET  
; CITY: WALTHAM  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/482,142  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/445,307  
; FILING DATE: 07 June 1995  
; ATTORNEY/AGENT INFORMATION:

; NAME: CRAIG, ANNE I.  
; REGISTRATION NUMBER: 32,976  
; REFERENCE/DOCKET NUMBER: 017.60US  
; TELEPHONE: (617) 466-6000  
; TELEFAX: (617) 466-6040  
; INFORMATION FOR SEQ ID NO: 206:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 29 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: internal  
; US-08-482-142-206

Query Match 42.5%; Score 77; DB 2; Length 29;  
Best Local Similarity 92.9%; Pred. No. 0.00035;  
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 19 HGSEPCNIHRGKPF 32  
||||| |||||  
Db 1 HGSEPCNIHRGKPF 14

RESULT 11  
US-08-478-572-43  
; Sequence 43, Application US/08478572  
; Patent No. 5968526  
; GENERAL INFORMATION:  
; APPLICANT: Garman, Richard  
; APPLICANT: Greenstein, Julia  
; APPLICANT: Kuo, Mei-chang  
; APPLICANT: Rogers, Bruce  
; APPLICANT: Franzen, Henry  
; APPLICANT: Chen, Xian  
; APPLICANT: Evans, Sean  
; APPLICANT: Shaked, Ze'ev  
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
; NUMBER OF SEQUENCES: 207  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
; STREET: 610 LINCOLN STREET  
; CITY: WALTHAM  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/478,572  
; FILING DATE: 07-June-1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/445,307  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CRAIG, ANNE I.  
; REGISTRATION NUMBER: 32,976  
; REFERENCE/DOCKET NUMBER: 017.60US  
; TELEPHONE: (617) 466-6000  
; TELEFAX: (617) 466-6040  
; INFORMATION FOR SEQ ID NO: 43:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 29 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide



```
;
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS
; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMUNOLOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,296
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: 08/445,307
; FILING DATE: 07 June 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAIG, ANNE I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: 017.605
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6040
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; US-08-484-296-43

Query Match 42.5%; Score 77; DB 4; Length 29;
Best Local Similarity 92.9%; Pred. No. 0.00035;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 19 HGSEPCNIHRGKPF 32
Db 1 HGSEPCIIHRGKPF 14

RESULT 15
US-08-484-296-119
; Sequence 119, Application US/08484296
; Patent No. 6268491
; GENERAL INFORMATION:
; APPLICANT: Garman, Richard
; APPLICANT: Greenstein, Julia
; APPLICANT: Kuo, Mei-chang
; APPLICANT: Rogers, Bruce
; APPLICANT: Franzen, Henry
; APPLICANT: Chen, Xian
; APPLICANT: Evans, Sean
; APPLICANT: Shaked, Ze'ev
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS
; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMUNOLOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,296
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: 08/445,307
; FILING DATE: 07 June 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAIG, ANNE I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: 017.605
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6040
; INFORMATION FOR SEQ ID NO: 119:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; US-08-484-296-119

Query Match 42.5%; Score 77; DB 4; Length 29;
Best Local Similarity 92.9%; Pred. No. 0.00035;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 19 HGSEPCNIHRGKPF 32
Db 1 HGSEPCIIHRGKPF 14

Search completed: December 4, 2002, 13:18:26
Job time : 15 secs
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 12:49:07 : Search time 6.07083 Seconds  
(without alignments)  
490.899 Million cell updates/sec

Title: US-09-362-731A-1  
Perfect score: 165  
Sequence: 1 QYIKANSKFCITELGGHEIKKLVPGCHGS 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84	50.9	146	2 A60381	allergen Der p II
2	75	45.5	1315	1 BTCLTN	tentoxilysin (EC 3
3	67	40.6	129	2 A61501	allergen Der f II
4	67	40.6	129	2 JU0394	allergen Der f II
5	67	40.6	138	2 B61241	allergen Der f II
6	67	40.6	138	2 A61241	allergen Der f II
7	55	33.3	487	1 S55194	DNA-directed DNA p
8	54	32.7	1333	2 S38635	blastopia polyprot
9	52	31.5	230	2 D84215	hypothetical prote
10	52	31.5	1386	2 S73401	MG064 homolog RO2
11	51	30.9	332	1 F69508	N-acetyl-gamma-glu
12	51	30.9	378	2 AF12180	hypothetical prote
13	50.5	30.6	244	2 S29982	class II histocomp
14	50.5	30.6	433	2 B82537	probable phosphodi
15	49	29.7	202	2 E54382	hypothetical prote
16	49	29.7	267	2 A64200	DNA polymerase III
17	49	29.7	394	2 F89780	ornithine aminotra
18	49	29.7	604	2 E75372	oligoendopeptidase
19	49	29.7	1882	2 S73484	hypothetical prote
20	48.5	29.4	167	2 C89606	protein B0416.2 [i
21	48.5	29.4	237	2 G75476	conserved hypothet
22	48	29.1	243	2 A60329	probable membrane
23	48	29.1	383	2 C69442	succinyl-CoA synth
24	48	29.1	440	1 XXRTN	phosphatidylcholin
25	48	29.1	511	2 E90083	t-complex protein1
26	48	29.1	4450	2 JX0340	gramicidin S synth
27	48	29.1	4452	1 YGBS62	gramicidin S synth
28	47.5	28.8	180	2 G86826	diamine N-acetyltr
29	47	28.5	130	2 AD1444	phage protein homo

30 47 28.5 147 2 A84546  
31 47 28.5 162 2 T49957  
32 47 28.5 188 2 PL0153  
33 47 28.5 189 2 G97978  
34 47 28.5 392 2 B69321  
35 47 28.5 438 1 XXMSN  
36 47 28.5 440 1 XXHUN  
37 47 28.5 440 2 JC1502  
38 47 28.5 458 2 T32634  
39 47 28.5 6486 2 T31076  
40 46.5 28.2 245 2 S29980  
41 46.5 28.2 1060 2 S06286  
42 46.5 28.2 1086 2 S16752  
43 46.5 28.2 1701 2 A54498  
44 46.5 28.2 1701 2 A26868  
45 46.5 28.2 1726 1 SAZQGM

#### ALIGNMENTS

##### RESULT 1

A60381  
allergen Der p II precursor - house-dust mite (Dermatophagoides pteronyssinus)  
C:Species: Dermatophagoides pteronyssinus  
C:Date: 03-Mar-1993 #sequence\_revision 03-Mar-1993 #text\_change 13-Sep-1998  
C:Accession: A60381  
R:Chua, K.Y.; Doyle, C.R.; Simpson, R.J.; Turner, K.J.; Stewart, G.A.; Thomas, W.R.  
Int. Arch. Allergy Appl. Immunol. 91, 118-123, 1990  
A:Title: Isolation of cDNA coding for the major mite allergen Der p II by IgE plaque  
A:Reference number: A60381; MUID:90256301; PMID:2341191  
A:Accession: A60381  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-146 <CHU>  
C:Superfamily: allergen Der p II  
F:1-17/Domain: signal sequence #status predicted <SIG>  
F:18-146/Product: allergen Der p II #status predicted <MAT>

Query Match 50.9%; Score 84; DB 2; Length 146;  
Best Local Similarity 63.6%; Pred. No. 0.00012;  
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKLVPGCHGS 31  
: : |||||  
Db 20 VDVKDCANHEIKKLVPGCHGS 41

##### RESULT 2

BTCLTN  
tentoxilysin (EC 3.4.24.68) precursor - Clostridium tetani  
N:Alternate names: tetanus neurotoxin  
C:Species: Clostridium tetani  
C:Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 03-Jun-2002  
C:Accession: A25689; A25194; B25194; A60759; S69348; S09364  
R:Eisel, U.; Jarausch, W.; Goretzki, K.; Henschen, A.; Engels, J.; Weller, U.; Hudel, E.MBO J. 5, 2495-2502, 1986  
A:Title: Tetanus toxin: primary structure, expression in E. coli, and homology with b  
A:Reference number: A25689; MUID:87053814; PMID:3536478  
A:Accession: A25689  
A:Molecule type: DNA  
A:Residues: 1-1315 <EIS>  
A:Cross-references: GB:X04436; NID:940769; PIDN:CAA28033.1; PID:g40770  
R:Fairweather, N.F.; Lyness, V.A.  
Nucleic Acids Res. 14, 7809-7812, 1986  
A:Title: The complete nucleotide sequence of tetanus toxin.  
A:Reference number: A25757; MUID:87040747; PMID:3774547  
A:Accession: A25757  
A:Molecule type: DNA  
A:Residues: 1-1315 <FAI>  
A:Cross-references: GB:X06214; NID:940773; PIDN:CAA29564.1; PID:g40774  
A:Experimental source: strain CN3911  
R:Fairweather, N.F.; Lyness, V.A.; Pickard, D.J.; Allen, G.; Thomson, R.O.

J. Bacteriol. 165, 21-27, 1986  
A:Title: Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C in E  
A:Reference number: A25194; MUID:86085672; PMID:3510187  
A:Accession: A25194  
A:Molecule type: DNA  
A:Residues: 743-1315 <FA2>  
A:Cross-references: GB:M12739; NID:g144920; PIDN:AAA23282.1; PID:g144921  
A:Accession: B25194  
A:Molecule type: protein  
A:Residues: 865-894 <FA3>  
R:Matsumoto, M.; Lei, D.L.; Sugimoto, N.; Ozutsumi, K.; Okabe, T.  
Infect. Immun. 57, 3588-3593, 1989  
A:Title: Isolation, purification, and characterization of fragment B, the NH-2-terminal  
A:Reference number: A60759; MUID:90035436; PMID:2478476  
A:Accession: A60759  
A:Molecule type: protein  
A:Residues: 461-475 <MAT>  
R:Demotz, S.; Lanzavecchia, L.; Eisel, U.; Niemann, H.; Widmann, C.; Corradin, G.  
J. Immunol. 142, 394-402, 1989  
A:Title: Delineation of several DR-restricted tetanus toxin T cell epitopes.  
A:Reference number: JS0098; MUID:89093918; PMID:2463305  
A:Contents: annotation; epitope region  
R:Schiavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; Dasgupta, B.  
Nature 359, 832-835, 1992  
A:Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteolyt  
A:Reference number: S27125; MUID:93063293; PMID:1331807  
A:Contents: annotation  
R:de Filippis, V.; Vangelista, L.; Schiavo, G.; Tonello, F.; Montecucco, C.  
Eur. J. Biochem. 229, 61-69, 1995  
A:Title: Structural studies on the zinc-endopeptidase light chain of tetanus neurotoxin.  
A:Reference number: S69348; MUID:95262688; PMID:7744050  
A:Accession: S69348  
A:Molecule type: protein  
A:Residues: 2-31 <DEF>  
C:Comment: The source of this protein was an extrachromosomal plasmid.  
C:Comment: The precursor is cleaved by endogenous proteinase activity to form light (fra  
dual chains are not toxic when separated). The amino end of the heavy chain (fragment B)  
C:Comment: Fragment B forms ion channels in a lipid bilayer. Fragment C binds to ganglio  
C:Comment: This potent neurotoxin binds to peripheral neuronal synapses, is internalized  
presynaptic neurons. It inhibits neurotransmitter release by proteolytic cleavage of sy  
C:Function:  
A:Description: blocks neuroexocytosis via hydrolysis of a Gln-Phe peptide bond in synapt  
C:Superfamily: tetanus toxin  
C:Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc  
F:2-457/Product: tetroxylisin light chain (fragment A) #status predicted <TTL>  
F:461-1315/Product: tetroxylisin heavy chain (fragment B C) #status experimental <TTH>  
F:461-864/Domain: channel forming (fragment B) #status predicted <TXB>  
F:865-1315/Domain: ganglioside binding (fragment C) #status predicted <TXC>  
F:233,237/Binding site: zinc (His) #status predicted  
F:234/Active site: Glu #status predicted

Query Match 45.5%; Score 75; DB 1; Length 1315;  
Best Local Similarity 81.0%; Pred. No. 0.023;  
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITELGGHEIK 21  
|||||  
Db 830 QYIKANSKFIGITELKKLESK 850

RESULT 3  
A61501  
allergen Der f II - house-dust mite (Dermatophagoides farinae) (fragment)  
C:Species: Dermatophagoides farinae  
C:Date: 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 13-Sep-1998  
C:Accession: A61501  
R:Trudinger, M.; Chua, K.Y.; Thomas, W.R.  
Clin. Exp. Allergy 21, 33-37, 1991  
A:Title: cDNA encoding the major mite allergen Der f II.  
A:Reference number: A61501; MUID:91215495; PMID:2021876  
A:Accession: A61501  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA

A:Residues: 1-129 <TRU>  
C:Superfamily: allergen Der p II

Query Match 40.6%; Score 67; DB 2; Length 129;  
Best Local Similarity 50.0%; Pred. No. 0.03;  
Matches 11; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31  
: : : : ||||| : |||||  
Db 3 VDVKDCANNEIKKVMVDGCHGS 24

RESULT 4  
JU0394  
allergen Der f II (pFL2) - house-dust mite (Dermatophagoides farinae)  
C:Species: Dermatophagoides farinae  
C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 17-Mar-1999  
C:Accession: JU0394  
R:Yuuki, T.; Okumura, Y.; Ando, T.; Yamakawa, H.; Suko, M.; Haida, M.; Okudaira, H.  
Agric. Biol. Chem. 55, 1233-1238, 1991  
A:Title: Cloning and expression of cDNA coding for the major house dust mite allergen  
A:Reference number: PS0417; MUID:91291341; PMID:1368682  
A:Accession: JU0394  
A:Molecule type: mRNA  
A:Residues: 1-129 <YUO>  
C:Superfamily: allergen Der p II

Query Match 40.6%; Score 67; DB 2; Length 129;  
Best Local Similarity 50.0%; Pred. No. 0.03;  
Matches 11; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31  
: : : : ||||| : |||||  
Db 3 VDVKDCANNEIKKVMVDGCHGS 24

RESULT 5  
B61241  
allergen Der f II precursor - house-dust mite (Dermatophagoides farinae) (fragment)  
C:Species: Dermatophagoides farinae  
C:Date: 12-May-1994 #sequence\_revision 27-Jun-1994 #text\_change 13-Sep-1998  
C:Accession: B61241; JU0395  
R:Yuuki, T.; Okumura, Y.; Ando, T.; Yamakawa, H.; Suko, M.; Haida, M.; Dohi, M.; Oku  
Int. Arch. Allergy Appl. Immunol. 94, 354-356, 1991  
A:Title: Synthesis of biologically active recombinant Der f II.  
A:Reference number: A61241; MUID:92040281; PMID:1937898  
A:Accession: B61241  
A:Molecule type: mRNA  
A:Residues: 1-138 <YUO>  
C:Superfamily: allergen Der p II  
F:1-9/Domain: signal sequence (fragment) #status predicted <SIG>  
F:10-138/Product: allergen Der f II #status predicted <MAT>

Query Match 40.6%; Score 67; DB 2; Length 138;  
Best Local Similarity 50.0%; Pred. No. 0.032;  
Matches 11; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31  
: : : : ||||| : |||||  
Db 12 VDVKDCANNEIKKVMVDGCHGS 33

RESULT 6  
A61241  
allergen Der f II precursor - house-dust mite (Dermatophagoides farinae) (fragment)  
C:Species: Dermatophagoides farinae  
C:Date: 12-May-1994 #sequence\_revision 27-Jun-1994 #text\_change 13-Sep-1998  
C:Accession: A61241; PS0417  
R:Yuuki, T.; Okumura, Y.; Ando, T.; Yamakawa, H.; Suko, M.; Haida, M.; Dohi, M.; Oku  
Int. Arch. Allergy Appl. Immunol. 94, 354-356, 1991  
A:Title: Synthesis of biologically active recombinant Der f II.  
A:Reference number: A61241; MUID:92040281; PMID:1937898  
A:Accession: A61241



[illegible]

Query Match		31.5%;	Score 52;	DB 2;	Length 1386;
Best Local Similarity		34.2%;	Pred. No. 51;		
Matches 13;		Conservative 7;	Mismatches 10;	Indels 8;	Gaps 1;
Qy	2	YIKANSKFIGITELGGHIKKV-----LVPGCHGS	31		
Db	538	YLKKNTEIGILKANGVSGRKINLSLLIFSLIPGIVGS	575		
RESULT 11					
F69508					
N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38) - Archaeoglobus fulgidus					
C:Species: Archaeoglobus fulgidus					
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999					
C:Accession: F69508					
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson					
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.					
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.					
Nature 390, 364-370, 1997					
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.					
Smith, H.O.; Woese, C.R.; Venter, J.C.					
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo					
A:Reference number: A69250; MUID:98049343; PMID:9389475					
A:Accession: F69508					
A:Status: preliminary; nucleic acid sequence not shown; translation not shown					
A:Molecule type: DNA					
A:Residues: 1-332 <KLE>					
C:Cross-references: GB:AE000961; GB:AE000782; NID:g2689284; PIDN:AAB89185.1; PID:g264846					
C:Superfamily: N-acetyl-gamma-glutamyl-phosphate reductase					
C:Keywords: oxidoreductase					
Query Match		30.9%;	Score 51;	DB 1;	Length 332;
Best Local Similarity		45.2%;	Pred. No. 16;		
Matches 14;		Conservative 4;	Mismatches 7;	Indels 6;	Gaps 2;
Qy	2	YIKANSKFIGITELGGHIKK---VLVPGCH	29		
Db	118	YVEA---VYGLTELHREIKKANLVANPGCY	145		
RESULT 12					
AI2180					
hypothetical protein all3000 [imported] - Nostoc sp. (strain PCC 7120)					
C:Species: Nostoc sp.					
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120					
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002					
C:Accession: AI2180					
R:kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuchih					
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S					
DNA Res. 8, 205-213, 2001					
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana					
A:Reference number: AB1807; MUID:21595285; PMID:11759840					
A:Accession: AI2180					
A:Status: preliminary					
A:Molecule type: DNA					
A:Residues: 1-378 <KUR>					
C:Cross-references: GB:BA000019; PIDN:BA874699.1; PID:g17132094; GSPDB:GN00179					
A:Experimental source: strain PCC 7120					
C:Genetics:					
A:Gene: all3000					
Query Match		30.9%;	Score 51;	DB 2;	Length 378;
Best Local Similarity		42.9%;	Pred. No. 18;		
Matches 9;		Conservative 6;	Mismatches 6;	Indels 0;	Gaps 0;
Qy	2	YIKANSKFIGITELGGHIKK	22		
Db	320	WFKSSKKYIGIIDVDGKEAKK	340		
RESULT 13					
S29982					
class II histocompatibility antigen - Atlantic salmon					

C:Species: Salmo salar (Atlantic salmon)	
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999	
C:Accession: S29982	
R:Hordvik, I.	
submitted to the EMBL Data Library, October 1992	
A:Reference number: S29980	
A:Accession: S29982	
A:Status: preliminary	
A:Molecule type: mRNA	
A:Residues: 1-244 <HOR>	
A:Cross-references: EMBL:X70166; NID:g64369; PID:g64370	
C:Superfamily: class II histocompatibility antigen; immunoglobulin homology	
Query Match	
Best Local Similarity	
Matches 11;	
Conservative 3;	
Mismatches 2;	
Indels 3;	
Gaps 1;	
Qy	1
Db	51
RESULT 14	
B82537	
probable phosphodiesterase I (EC 3.1.4.1) / nucleotide diphosphatase (EC 3.6.1.9) XF2	
C:Species: Xylella fastidiosa	
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 16-Aug-2002	
C:Accession: B82537	
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq	
Nature 406, 151-157, 2000	
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.	
A:Reference number: A82515; MUID:20365717; PMID:10910347	
A:Note: for a complete list of authors see reference number A59328 below	
A:Accession: B82537	
A:Molecule type: DNA	
A:Residues: 1-433 <SIM>	
A:Cross-references: GB:AE004067; GB:AE003849; NID:g9107818; PIDN:AAF85396.1; GSPDB:GN	
A:Experimental source: strain 9a5C	
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.	
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carfer	
as-Neto, E.; Docena, C.; El-dorri, H.; Facincani, A.P.; Ferreira, A.J.S.	
submitted to GenBank, June 2000	
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr	
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Klieger, J.E.; Kuramae, E.E.; La	
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins	
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Miyaki, C.	
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa	
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv	
M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.	
A:Reference number: A59328	
A:Contents: annotation	
C:Genetics:	
A:Gene: XF2599	
C:Superfamily: human phosphodiesterase I / nucleotide pyrophosphatase 4	
C:Keywords: coenzyme A; metalloprotein; phosphoric diester hydrolase	
Query Match	
Best Local Similarity	
Matches 10;	
Conservative 10;	
Mismatches 10;	
Indels 1;	
Gaps 1;	
Qy	2
Db	273
RESULT 15	
E64362	
hypothetical protein MJ0501 - Methanococcus jannaschii	
C:Species: Methanococcus jannaschii	
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000	
C:Accession: E64362	
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak	

; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;  
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
Science 273, 1058-1073, 1996  
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.  
A:Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*  
A:Reference number: A64300; MUID:96337999; PMID:8688087  
A:Accession: E54362  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-202 <BUIL>  
A:Cross-references: GB:U67500; GB:L77117; NID:g1591202; PIDN:AAB98492.1; PID:g1591204; T  
C:Genetics:  
A:Map position: REV443001-442393  
A:Start codon: GTG  
C:Superfamily: hypothetical protein MJ0501

Query Match 29.7%; Score 49; DB 2; Length 202;  
Best Local Similarity 37.5%; Pred. No. 19;  
Matches 9; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY      8 KFIGITELGGHEIKKVLVPGCHGS 31  
         | : | | | : : | : | :  
Db      2 KIMEIFEFGNGVKLFIGGLHGN 25

Search completed: December 4, 2002, 12:54:21  
Job time : 8.07083 secs

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 12:49:06 ; Search time 3.35833 Seconds  
(without alignments)  
382.858 Million cell updates/sec

Title: US-09-362-731a-1

Perfect score: 165

Sequence: 1 QYIKANSKFIGITELGGHEIKKVLPGCHGS 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84	50.9	146	1 ALL2_DERPT	P49278 dermatophag
2	75	45.5	1314	1 TETX_CLOTE	P04558 clostridium
3	74	44.8	145	1 ALL2_BURMA	Q04222 euroglyphus
4	67	40.6	146	1 ALL2_DERFA	Q00855 dermatophag
5	55	33.3	487	1 DP02_YEAST	P46957 saccharomyc
6	52	31.5	1386	1 Y064_MYCPN	P75613 mycoplasma
7	51	30.9	332	1 ARGC_ARCFU	Q28208 archaeoglob
8	50.5	30.6	347	1 ARGC_CORGL	Q59279 corynebacte
9	49	29.7	202	1 Y501_WETJA	Q57924 methanococc
10	49	29.7	259	1 LPXA_CHRVI	Q46481 chromatium
11	49	29.7	364	1 DP3B_MYCGE	P47247 mycoplasma
12	49	29.7	1882	1 Y468_MYCPN	P75109 mycoplasma
13	48.5	29.4	167	1 Y742_CAEEL	Q11070 caenorhabdi
14	48	29.1	382	1 SUC1_ARCFU	Q28732 archaeoglob
15	48	29.1	440	1 LCAT_RAT	P18424 rattus norv
16	48	29.1	4451	1 GRSB_BACBR	P14688 b gramicidi
17	47	28.5	188	1 LCAT_PIG	Q30930 sus scrofa
18	47	28.5	392	1 FT22_ARCFU	Q29685 archaeoglob
19	47	28.5	438	1 LCAT_MOUSE	P16301 mus musculu
20	47	28.5	440	1 LCAT_HUMAN	Q04180 homo sapien
21	47	28.5	440	1 LCAT_PAPAN	Q08758 papio anubi
22	47	28.5	440	1 LCAT_RABIT	P53761 oryctolagus
23	47	28.5	6486	1 TYCC_BACBR	Q30409 b tyrocidin
24	46.5	28.2	207	1 REGQ_BPPS3	O80286 bacterioph
25	46.5	28.2	1682	1 MSP1_PLAF3	P19598 plasmodium
26	46.5	28.2	1701	1 MSP1_PLAF6	P13819 plasmodium
27	46.5	28.2	1701	1 MSP1_PLAFM	P08569 plasmodium
28	46.5	28.2	1726	1 MSP1_PLAFC	P04934 plasmodium
29	46.5	28.2	1726	1 MSP1_PLAFD	P50495 plasmodium
30	46	27.9	196	1 PTH_YERPE	Q82ey4 yersinia pe
31	46	27.9	239	1 SPSA_PYRAE	Q82t87 pyrobaculum
32	46	27.9	337	1 MALR_STAXY	Q56201 staphylococ
33	46	27.9	424	1 OAT_YEAST	P07991 saccharomyc

## ALIGNMENTS

### RESULT 1

ID	ALL2_DERPT	STANDARD;	PRT;	146 AA.
AC	P49278;			
DT	01-FEB-1996 (Rel. 33, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Mite group 2 allergen Der p 2 precursor (Der p II) (DPX).			
GN	DERP2.			
OS	Dermatophagoides pteronyssinus (House-dust mite).			
OC	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;			
OC	Acariformes; Sarcophtiformes; Astigmata; Analgoidea; Pyroglyphidae;			
OC	Dermatophagoides.			
OX	NCBI_TaxID=6956;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90256301; PubMed=2341191;			
RA	Chua K.Y., Doyle C.R., Simpson R.J., Turner K.J., Stewart G.A.,			
RA	Thomas W.R.;			
RT	"Isolation of cDNA coding for the major mite allergen Der p II by IgE			
RT	plaque immunoassay.";			
RL	Int. Arch. Allergy Appl. Immunol. 91:118-123(1990).			
[2]				
RP	SEQUENCE FROM N.A., AND VARIANTS.			
RX	MEDLINE=21290932; PubMed=11398075;			
RA	Smith W.-A., Hales B.J., Jarnicki A.G., Thomas W.R.;			
RT	"Allergens of wild house dust mites: environmental Der p 1 and Der p 2			
RT	sequence polymorphisms.";			
RL	J. Allergy Clin. Immunol. 107:985-992(2001).			
[3]				
RP	PARTIAL SEQUENCE OF 18-57.			
RX	MEDLINE=89278484; PubMed=2732406;			
RA	Heymann P.W., Chapman M.D., Aalberse R.C., Fox J.W.,			
RA	Platts-Mills T.A.;			
RT	"Antigenic and structural analysis of group II allergens (Der f II			
RT	and Der p II) from house dust mites (Dermatophagoides spp).";			
RL	J. Allergy Clin. Immunol. 83:1055-1067(1989).			
[4]				
RP	STRUCTURE BY NMR.			
RX	MEDLINE=98409423; PubMed=9737847;			
RA	Mueller G.A., Benjamin D.C., Rule G.S.;			
RT	"Tertiary structure of the major house dust mite allergen Der p 2:			
RT	sequential and structural homologues.";			
RL	Biochemistry 37:12707-12714(1998).			
CC	-I- SUBCELLULAR LOCATION: Secreted.			
CC	-I- SIMILARITY: BELONGS TO THE NPC2 FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; AF276239; AAF86462.1; -.			

P57693 thermoplasma  
P49847 drosophila  
Q00971 vibrio prot  
P56002 helicobacte  
P47310 mycoplasma  
P26534 pylaiella 1  
P12685 saccharomyc  
P28569 saccharomyc  
Q9cmz1 pasteurella  
O19885 cyanidium c  
Q9xek6 tortula rur  
O02380 tyrophagus

34 46 27.9 499 1 SYFA\_THEAC  
35 46 27.9 592 1 T2D5\_DROME  
36 46 27.9 609 1 NPROV\_VIBPR  
37 46 27.9 691 1 EFG\_HELPY  
38 46 27.9 1331 1 Y064\_MYCGE  
39 45.5 27.6 132 1 TRPE\_PYLIT  
40 45.5 27.6 1235 1 TRKL\_YEAST  
41 45.5 27.6 1241 1 TRKL\_SACBA  
42 45.5 27.6 1905 1 Y659\_PASMU  
43 45 27.3 100 1 RK27\_CYACA  
44 45 27.3 134 1 RS14\_TORRU  
45 45 27.3 141 1 ALL2\_TYRPU

```

DR PDB; 1A9V; 14-OCT-98.
DR InterPro; IPR003172; E1_DerP2_DerF2.
DR Pfam; PF02221; E1_DerP2_DerF2; 1.
KW Allergen; Signal; 3D-structure; Polymorphism.
FT SIGNAL 1 17 MITE GROUP 2 ALLERGEN DER P 2.
FT CHAIN 18 146
FT DISULFID 25 136
FT DISULFID 38 44
FT DISULFID 90 95
FT VARIANT 39 39
FT VARIANT 40 40
FT VARIANT 44 44
FT VARIANT 47 47
FT VARIANT 49 49
FT VARIANT 56 56
FT VARIANT 57 57
FT VARIANT 61 61
FT VARIANT 64 64
FT VARIANT 75 75
FT VARIANT 78 78
FT VARIANT 81 81
FT VARIANT 95 95
FT VARIANT 98 98
FT VARIANT 108 108
FT VARIANT 111 111
FT VARIANT 114 114
FT VARIANT 115 115
FT VARIANT 116 116
FT VARIANT 118 118
FT VARIANT 127 127
FT VARIANT 128 128
FT VARIANT 131 131
FT VARIANT 133 133
FT VARIANT 144 144
SQ SEQUENCE 146 AA; 15999 MW; 591B2FA7FD26D3AF CRC64;

Query Match      50.9%; Score 84; DB 1; Length 146;
Best Local Similarity 63.6%; Pred. No. 2.3e-05;
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31
   : : | | | | | | | | | |
Db 20 VDVKDCANHEIKKVLVPGCHGS 41

RESULT 2
TETX_CLOTE
ID TETX_CLOTE STANDARD; PRT; 1314 AA.
AC P04958;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Tetanus toxin precursor (EC 3.4.24.68) (Tentoxylisin).
OS Clostridium tetani.
OG Plasmid.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1513;
RX MEDLINE=87053814; PubMed=3536478;
RA Eisel U., Jarausch W., Goratzki K., Henschen A., Engels J.,
RT Weller U., Hudei M., Habermann E., Niemann H.;
RT "Tetanus toxin: primary structure, expression in E. coli, and
RT homology with botulinum toxins.";
RL EMBO J. 5:2495-2502(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CN3911;
RX MEDLINE=87040747; PubMed=3774547;
RA Fairweather N.F., Lyness V.A.;
RT "The complete nucleotide sequence of tetanus toxin.";
RL Nucleic Acids Res. 14:7809-7812(1986).

RN [3]
RP SEQUENCE OF 742-1314 FROM N.A.
RX MEDLINE=86085672; PubMed=3510187;
RA Fairweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.;
RT "Cloning, nucleotide sequencing, and expression of tetanus toxin
RT fragment C in Escherichia coli.";
RL J. Bacteriol. 165:21-27(1986).
RN [4]
RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=90201034; PubMed=2108021;
RA Kriegelstein K., Henschen A., Weller U., Habermann E.;
RT "Arrangement of disulfide bridges and positions of sulphydryl groups
RT in tetanus toxin.";
RL Eur. J. Biochem. 188:39-45(1990).
RN [5]
RP PARTIAL SEQUENCE.
RX MEDLINE=92037649; PubMed=1935979;
RA Kriegelstein K.G., Henschen A.H., Weller U., Habermann E.;
RT "Limited proteolysis of tetanus toxin. Relation to activity and
RT identification of cleavage sites.";
RL Eur. J. Biochem. 202:41-51(1991).
RN [6]
RP IDENTIFICATION AS ZINC-PROTEASE.
RX MEDLINE=93010948; PubMed=1396558;
RA Schiavo G., Poulain B., Rossetto O., Benfenati F., Tauc L.,
RA Montecucco C.;
RT "Tetanus toxin is a zinc protein and its inhibition of
RT neurotransmitter release and protease activity depend on zinc.";
RL EMBO J. 11:3577-3583(1992).
RN [7]
RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=93063293; PubMed=1331807;
RA Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,
RA Dasgupta B.R., Montecucco C.;
RT "Tetanus and botulinum-B neurotoxins block neurotransmitter release
RT by proteolytic cleavage of synaptobrevin.";
RL Nature 359:832-835(1992).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 874-1314.
RX MEDLINE=97475217; PubMed=9334741;
RA Umland T.C., Wingert L.M., Swaminathan S., Furey W.F., Schmidt J.J.,
RA Sax M.;
RT "Structure of the receptor binding fragment HC of tetanus
RT neurotoxin.";
RL Nat. Struct. Biol. 4:788-792(1997).
CC -!- FUNCTION: TETANUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 76-GLN-1-PHE-77
CC BOND OF SYNAPTOSOMAL-2.
CC -!- CATALYTIC ACTIVITY: HYDROLYSIS OF 76-GLN-1-PHE-77 BOND IN
CC SYNAPTOSOMAL-2.
CC -!- SUBUNIT: THE PRECURSOR POLYPEPTIDE IS SUBSEQUENTLY CLEAVED TO
CC YIELD SUBCHAINS L AND H. THESE REMAIN LINKED BY A DISULFIDE BRIDGE
CC AND ARE NON-TOXIC AFTER SEPARATION.
CC -!- MISCELLANEOUS: THE C-TERMINAL OF THE HEAVY CHAIN BINDS TO
CC GANGLIOSIDE RECEPTORS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M7.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X04436; CAA28033.1; -
CC DR EMBL; M12739; AAA23282.1; -
CC DR EMBL; X06214; CAA29564.1; -
CC DR PIR; A25689; BTCLTN.

```

```

DR  PDB; 1AF9; 29-APR-98.
DR  PDB; 1A8D; 14-OCT-98.
DR  MEROPS; M27.001; -.
DR  InterPro; IPR000395; Bontoxilysin.
DR  InterPro; IPR000130; Zn_Mrptcdase.
DR  Pfam; PF01742; Peptidase_M27; 1.
DR  PRINTS; PR00760; BONTOXILYSIN.
DR  ProDom; PD001963; Bontoxilysin; 1.
DR  ProSITE; PS00142; ZINC_PROTEASE; 1.
KW  Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc; Plasmid;
KW  3D-structure.
FT  INIT_MET 0
FT  CHAIN 1 456
FT  CHAIN 457 1314
FT  METAL 232 232
FT  ACT_SITE 233 233
FT  METAL 236 236
FT  TRANSMEM 226 246
FT  TRANSMEM 669 689
FT  DISULFID 438 466
FT  DISULFID 1076 1092
SQ  SEQUENCE 1314 AA; 150550 MW; 134C3657133EF9D1D CRC64;

Query Match 45.5%; Score 75; DB 1; Length 1314;
Best Local Similarity 81.0%; Pred. No. 0.0047;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 QYIKANSKFGITELGSGHEIK 21
Db .829 QYIKANSKFGITELGSGHEIK 849

RESULT 3
ALL2_EURMA STANDARD; PRT; 145 AA.
AC Q9TZZ2; Q96430;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mite group 2 allergen Eur m 2 precursor (Eur m 2.0101/Eur m 2.0102).
GN EURM2.
OS Eukaryophus maynei (House-dust mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcophormes; Astigmata; Analgoidea; Pyroglyphidae;
OC Eukaryophus.
OX NCBI_TaxID=6958;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99126275; PubMed=9925958;
RA Smith W., Mills K., Hazell L., Hart B.J., Thomas W.;
RT "Molecular analysis of the group 1 and 2 allergens from the house dust
Int. Arch. Allergy Immunol. 118:15-22(1999).
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- SIMILARITY: BELONGS TO THE NPC2 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF047613; AAC82349.1; -
CC EMBL; AF047614; AAC82350.1; -
CC HSP; P43278; IA9V.
CC InterPro; IPR003172; El_DerP2_DerF2.
CC Pfam; PF02221; El_DerP2_DerF2; 1.
KW Allergen; Signal; Polymorphism.
FT SIGNAL 1 16
FT CHAIN 17 145
FT DISULFID 24 135
FT SIGNAL 17 145
FT CHAIN 17 145
FT DISULFID 24 135

```

```

FT DISULFID 37 43
FT DISULFID 89 94
FT VARIANT 21 21
SQ SEQUENCE 145 AA; 15747 MW; 6655B16C8503A565 CRC64;

Query Match 44.8%; Score 74; DB 1; Length 145;
Best Local Similarity 59.1%; Pred. No. 0.00069;
Matches 13; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31
Db 19 VDIKCANHEIKKVMVPGCKGS 40

RESULT 4
ALL2_DERFA STANDARD; PRT; 146 AA.
AC Q00855; P39672; Q26359;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mite group 2 allergen Der f 2 precursor (Der f II).
GN DERF2.
OS Dermatophagoides farinae (House-dust mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcophormes; Astigmata; Analgoidea; Pyroglyphidae;
OC Dermatophagoides.
OX NCBI_TaxID=6954;
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RX MEDLINE=91291341; PubMed=1368682;
RA Yuuki T., Okumura Y., Ando T., Yamakawa H., Suko M., Haida M.,
RT "Cloning and expression of cDNA coding for the major house dust mite
allergen Der f II in Escherichia coli.";
RL Agric. Biol. Chem. 55:1233-1238(1991).
RN [2]
RP SEQUENCE OF 4-146 FROM N.A.
RX MEDLINE=94256850; PubMed=8198452;
RA Okuhira H.;
RT "Molecular biology of mite antigens.";
RL Arerugi 43:435-440(1994).
RN [3]
RP DISULFIDE BONDS, AND PARTIAL SEQUENCE.
RX MEDLINE=93283958; PubMed=8508052;
RA Nishiyama C., Yuuki T., Takai T., Okumura Y., Okudaira H.;
RT "Determination of three disulfide bonds in a major house dust mite
allergen, Der f II.";
RL Int. Arch. Allergy Immunol. 101:159-166(1993).
RN [4]
RP PARTIAL SEQUENCE OF 18-52.
RX MEDLINE=89278484; PubMed=2732406;
RA Heymann P.W., Chapman M.D., Aalberse R.C., Fox J.W.,
RA Platts-Mills T.A.;
RT "Antigenic and structural analysis of group II allergens (Der f II
and Der p II) from house dust mites (Dermatophagoides spp).";
RL J. Allergy Clin. Immunol. 83:1055-1067(1989).
RN [5]
RP STRUCTURE BY NMR.
RX MEDLINE=98079068; PubMed=9417088;
RA Ichikawa S., Hatanaka H., Yuuki T., Iwamoto N., Kojima S.,
RA Nishiyama C., Ogura K., Okumura Y., Inagaki F.;
RT "Solution structure of Der f 2, the major mite allergen for atopic
diseases.";
RL J. Biol. Chem. 273:356-360(1998).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- MISCELLANEOUS: THE SEQUENCE SHOWN HERE IS FROM CLONE 2. THE
CC N-TERMINAL SEQUENCE (AA 1-8) FROM CLONE 1 AND 11 ARE NOT YET
CC KNOWN.
CC -!- SIMILARITY: BELONGS TO THE NPC2 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D10447; BAA01239.1; -.
DR EMBL; D10448; BAA01240.1; -.
DR EMBL; D10449; BAA01241.1; -.
DR EMBL; S70378; AAB30829.1; -.
DR PIR; PS0417; PS0417.
DR PDB; IAHK; 08-APR-98.
DR PDB; IAHM; 08-APR-98.
DR InterPro; IPR003172; E1_DerP2_DerF2.
DR Pfam; PF02221; E1_DerP2_DerF2; 1.
KW Allergen; Signal; Polymorphism; 3D-structure.
FT SIGNAL 1 17
FT CHAIN 18 146 MITE GROUP 2 ALLERGEN DER F 2.
FT DISULFID 25 136
FT DISULFID 38 44.
FT DISULFID 90 95
FT VARIANT 93 93 M -> V (IN CLONE 1).
FT VARIANT 105 105 I -> A (IN CLONE 11).
FT VARIANT 108 108 I -> V (IN CLONE 11).
FT VARIANT 128 128 G -> A (IN CLONE 11).
FT VARIANT 142 142 ILCL -> GTMV (IN REF. 2).
FT CONFLICT 5 8
SQ SEQUENCE 146 AA; 15802 MW; FAL18206CD88534A CRC64;

Query Match 40.6%; Score 67; DB 1; Length 146;
Best Local Similarity 50.0%; Pred. No. 0.0077;
Matches 11; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31
: : : : :| | | | | | | | | |
DB 20 VDVKCANNEIKKVVWDGCHGS 41

RESULT 5
DPD2_YEAST
ID DPD2_YEAST STANDARD; PRT; 487 AA.
AC P46957;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA polymerase delta small subunit (EC 2.7.7.7).
GN POL31 OR HYS2 OR HUS2 OR SDP5 OR YJR006W OR J1427 OR YJR83.7.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Sugimoto K., Sakamoto Y., Matsumoto K.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA de Haan M., Smits P.H.M., Grivell L.A.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP PARTIAL SEQUENCE, AND CHARACTERIZATION.
RX MEDLINE=98083176; PubMed=9421503;
RA Hashimoto K., Nakashima N., Ohara T., Maki S., Sugino A.;
RT "The second subunit of DNA polymerase III (delta) is encoded by the
RT HYS2 gene in Saccharomyces cerevisiae."
RL Nucleic Acids Res. 26:477-485(1998).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=98344072; PubMed=9677405;
RA Gerik K.J., Li X., Pautz A., Burgers P.M.;
RT "Characterization of the two small subunits of Saccharomyces
RT cerevisiae DNA polymerase delta."
RL J. Biol. Chem. 273:19747-19755(1998).

```

```

CC -!- FUNCTION: REQUIRED FOR REPLICATION OF THE LEADING DNA STRAND AND
CC FOR COMPLETION OF LAGGING STRAND SYNTHESIS.
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + [DNA](N).
CC -!- SUBUNIT: HETEROTRIMER WITH SUBUNITS OF 125 KDA, 58 KDA AND 55 KDA.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: BELONGS TO THE DNA POLYMERASE DELTA/II SMALL SUBUNIT
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D50324; BAA08859.1; -.
DR EMBL; X87611; CAA60928.1; -.
DR EMBL; Z49506; CAA89528.1; -.
DR SGD; S0003766; HYS2.
DR Transferase; DNA-directed DNA polymerase; DNA replication;
KW Nuclear protein.
KW CONFLICT 156 156 L -> H (IN REF. 1).
FT CONFLICT 465 465 S -> N (IN REF. 1).
SQ SEQUENCE 487 AA; 55296 MW; F9E200BFE97A2C07 CRC64;

Query Match 33.3%; Score 55; DB 1; Length 487;
Best Local Similarity 44.4%; Pred. No. 1.6;
Matches 12; Conservative 4; Mismatches 9; Indels 2; Gaps 1;

QY 2 YIKANSKFTG--ITELGGHEIKKVLVP 26
| | | | | : | | | | | | | |
DB 436 YIVANQYFGRVVEIGGKNIKIISVP 462

RESULT 6
Y064_MYCPN
ID Y064_MYCPN STANDARD; PRT; 1386 AA.
AC P75613;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MG064 homolog (R02_orf1386).
GN MPN080 OR MP075.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelsreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae."
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: SOME, TO MG468.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF000009; AAB95723.1; -.
DR InterPro; IPR003838; DUF214.
DR Pfam; PF02687; DUF214; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 17 37 POTENTIAL.

```

DR PROSITE; PS01224; ARGC; 1.  
KW Arginine biosynthesis; Oxidoreductase; NADP; Complete proteome.  
FT ACT\_SITE 144 BY SIMILARITY.  
SQ SEQUENCE 332 AA; 37304 MW; 295EB7504EECE33 CRC64;

Query Match 30.9%; Score 51; DB 1; Length 332;  
Best Local Similarity 45.2%; Pred. No. 4.3;  
Matches 14; Conservative 4; Mismatches 7; Indels 6; Gaps 2;

QY 2 YIRANSKFIGITELGGHEIKK--VLVPGCH 29  
::| | :||| |||| | |||:  
Db 118 YVEA---VYLTELHREIRKANLVANPGCY 145

RESULT 8  
ARGC\_CORGL STANDARD; PRT; 347 AA.

ID ARGC\_CORGL STANDARD; PRT; 347 AA.  
AC Q59279; O32353;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38) (AGPR) (N-  
acetyl-glutamate semialdehyde dehydrogenase) (NAGSA dehydrogenase).  
GN ARGC OR CGLI394.  
OS Corynebacterium glutamicum (Brevibacterium flavum).  
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Corynebacteriaceae;  
OC Corynebacterium.  
OX NCBI\_TaxID=1718;  
RN [1]  
RS SEQUENCE FROM N.A.  
RC STRAIN=ATCC 13059 / AS019;  
RL Chun J.Y., Lee E.J., Cheon C.I., Min K.H., Lee M.-S.;  
RR Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.  
RN [2]  
RS SEQUENCE FROM N.A.  
RC STRAIN=ATCC 13059 / AS019;  
RL Park M.Y., Chun J.Y., Ko S.-Y., Lee M.-S.;  
RR "Molecular cloning of the arginine biosynthetic genes from  
RT Corynebacterium glutamicum";  
RN [3]  
RS Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.  
RN [4]  
RS SEQUENCE FROM N.A.  
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;  
RL Nakagawa S.;  
RR "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";  
RN [5]  
RS Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.  
RN [6]  
RS SEQUENCE OF 243-347 FROM N.A.  
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;  
RX MEDLINE=96146054; PubMed=8581175;  
RA Sakanyan V., Petrosyan P., Lecoq M., Boyen A., Legrain C.,  
RM Demarez M.N., Hallet J.N., Glansdorff N.;  
RT "Genes and enzymes of the acetyl cycle of arginine biosynthesis in  
RT Corynebacterium glutamicum: enzyme evolution in the early steps of  
RT the arginine pathway.";  
RN Microbiology 142:99-108(1996).  
CC -!- CATALYTIC ACTIVITY: N-acetyl-L-glutamate 5-semialdehyde + NADP(+)  
+ phosphate = N-acetyl-5-glutamyl phosphate + NADPH.  
CC -!- PATHWAY: Arginine biosynthesis; third step.  
CC -!- SIMILARITY: BELONGS TO THE NAGSA DEHYDROGENASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).

EMBL; AF005242; AAC62245.1; -  
DR EMBL; AF049897; AAC24812.1; -  
DR EMBL; AP005278; BAB98787.1; -





```

RP RP
RC SEQUENCE FROM N.A.
RX STRAIN=ATCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fleischmann J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Uterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Brott K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RA "The minimal gene complement of Mycoplasma genitalium.";
RT Science 270:397-403(1995).
RL [2]
RP REVISIONS.
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Uterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Brott K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RA Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RL [3]
RN SEQUENCE OF 267-364 FROM N.A.
RP STRAIN=ATCC 33530 / G-37;
RC MEDLINE=94364962; PubMed=8083173;
RX Bailey C.C., Bott K.F.;
RA "An unusual gene containing a dnaJ N-terminal box flanks the putative
RT origin of replication of Mycoplasma genitalium.";
RL J. Bacteriol. 176:5814-5819(1994).
CC -!- FUNCTION: DNA POLYMERASE III IS A COMPLEX, MULTICHAIN ENZYME
CC RESPONSIBLE FOR MOST OF THE REPLICATIVE SYNTHESIS IN BACTERIA.
CC THIS DNA POLYMERASE ALSO EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.
CC THE BETA CHAIN IS REQUIRED FOR INITIATION OF REPLICATION ONCE IT
CC IS CLAMPED ONTO DNA, IT SLIDES FREELY (BIDIRECTIONAL AND ATP-
CC INDEPENDENT) ALONG DUPLICATION DNA (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + [DNA](N).
CC -!- SUBUNIT: CONTAINS A CORE (COMPOSED OF ALPHA, EPSILON, AND
CC THETA CHAINS) THAT CAN REPAIR SHORT GAPS CREATED BY NUCLEASE
CC IN DUPLICATION. FOR EFFICIENT REPLICATION OF THE LONG, SINGLE-
CC STRANDED TEMPLATES, POL III REQUIRES THE AUXILIARY CHAINS BETA,
CC GAMMA, AND DELTA (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U39679; AAC71217.1; -.
CC EMBL; U09251; AAA57069.1; -.
CC TIGR; MG001; -.
CC InterPro; IPR001001; DNA_polIII_beta.
CC Pfam; PF00712; DNA_pol3_beta; 1.
CC Pfam; PF02768; DNA_pol3_beta_3; 1.
CC SMART; SM00480; POL3bc; 1.
CC TIGRFAMs; TIGR00663; dnai; 1.
CC Transferase; DNA-directed DNA polymerase; DNA replication;
CC Complete proteome.
CC SEQUENCE 364 AA; 42399 MW; 4D8C23EF9BC48624 CRC64;
DR DR
DR EMBL; U39679; AAC71217.1; -.
DR EMBL; U09251; AAA57069.1; -.
DR TIGR; MG001; -.
DR InterPro; IPR001001; DNA_polIII_beta.
DR Pfam; PF00712; DNA_pol3_beta; 1.
DR Pfam; PF02768; DNA_pol3_beta_3; 1.
DR SMART; SM00480; POL3bc; 1.
DR TIGRFAMs; TIGR00663; dnai; 1.
KW Transferase; DNA-directed DNA polymerase; DNA replication;
KW Complete proteome.
SQ SEQUENCE 364 AA; 42399 MW; 4D8C23EF9BC48624 CRC64;

Query Match 29.7%; Score 49; DB 1; Length 364;
Best Local Similarity 43.5%; Pred. No. 9.4;
Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 4 KANSKFIGITELGGHEIKKVLVP 26
: III:
: III:
Db 340 QGNSKYFLITSKSEPELQILVP 362

RESULT 12
Y468_MYCPN STANDARD; PRT; 1882 AA.
ID P75109; Q50317;
AC 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protein MG468 homolog (K05_orf1882).
GN MPN684 OR MP158.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [2]
RP SEQUENCE OF 1-1848 FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=96177562; PubMed=8604303;
RA Hilbert H., Himmelreich R., Plagens H., Herrmann R.;
RT "Sequence analysis of 56 kb from the genome of the bacterium Mycoplasma pneumoniae comprising the dnaA region, the atp operon and a
RT cluster of ribosomal protein genes.";
RL Nucleic Acids Res. 24:628-639(1996).
RN [3]
RP IDENTIFICATION BY MASS SPECTROMETRY.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=21088919; PubMed=11271496;
RA Regula J.T., Ueberle B., Boguth G., Goerg A., Schnoelzer M.,
RA Herrmann R., Frank R.;
RT "Towards a two-dimensional proteome map of Mycoplasma pneumoniae.";
RL Electrophoresis 21:3765-3780(2000).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: SOME, TO MG064.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000017; AAB95806.1; -.
CC EMBL; U34816; AAC43650.1; -.
CC Transmembrane; Complete proteome.
CC TRANSMEM 16 36 POTENTIAL.
CC TRANSMEM 987 1007 POTENTIAL.
CC TRANSMEM 1037 1057 POTENTIAL.
CC TRANSMEM 1080 1100 POTENTIAL.
CC TRANSMEM 1154 1174 POTENTIAL.
CC TRANSMEM 1759 1779 POTENTIAL.
CC TRANSMEM 1807 1827 POTENTIAL.
CC TRANSMEM 1828 1848 POTENTIAL.
CC TRANSMEM 1851 1871 POTENTIAL.
CC SEQUENCE 1882 AA; 209442 MW; 03CFA4D999A7120ED CRC64;

Query Match 29.7%; Score 49; DB 1; Length 1882;
Best Local Similarity 31.8%; Pred. No. 50;
Matches 7; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

QY 3 IKANSKFIGITELGGHEIKKVL 24
: III:
: III:
Db 1457 ISPSKFNLTDSGNSLSRLI 1478

RESULT 13
```

```

YT42_CABEL
ID YT42_CABEL STANDARD; PRT; 167 AA.
AC Q11070;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 19.0 kDa protein B0416.2 in chromosome X precursor.
GN B0416.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Favella T.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: STRONG, TO C.ELEGANS ZK856.6.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U23516; AAC38879.1; -.
DR WormPep; B0416.2; CE02432.
KW Hypothetical protein; Signal.
FT SIGNAL 1 25
FT CHAIN 26 167 HYPOTHETICAL PROTEIN B0416.2.
SQ SEQUENCE 167 AA; 19026 MW; CAFB029AC3CFD106 CRC64;
-----
Query Match 29.4%; Score 48.5; DB 1; Length 167;
Best Local Similarity 31.2%; Pred. No. 5;
Matches 10; Conservative 3; Mismatches 12; Indels 7; Gaps 1;

QY 6 NSKFIGITELGGHEIKKVLV-----PCCHG 30
| : : : | : | : |
Db 74 NGSLEILQASGHKFSKTHVQCLEDRPSCG 105
| : : : | : | : |

RESULT 14
SUC1_ARCFU
ID SUC1_ARCFU STANDARD; PRT; 382 AA.
AC O28732;
DT 30-MAY-2000 (Rel. 39, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Succinyl-CoA synthetase beta chain 1 (EC 6.2.1.5) (SCS-beta 1).
GN SUC1 OR AF1540.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Wenter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";

```

```

RL Nature 390:364-370(1997).
CC -!- CATALYTIC ACTIVITY: ATP + succinate + CoA = ADP + succinyl-CoA +
CC phosphate.
CC -!- PATHWAY: Tricarboxylic acid cycle.
CC -!- SUBUNIT: Composed of an alpha chain and a beta chain (By
CC similarity).
CC -!- SIMILARITY: BELONGS TO THE SUCCINYL-COA SYNTHETASE BETA SUBUNIT
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000996; AAB89706.1; ALT_INIT.
DR HSSP; P07460; 2SCU.
DR TIGR; AF1540; -.
DR InterPro; IPR0031135; ATP-grasp.
DR InterPro; IPR000303; CoA_ligase.
DR Pfam; PF00349; ligase-CoA; 1.
DR Pfam; PF02222; ATP-grasp; 1.
DR TIGRFAMs; TIGR01016; sucCoABeta; 1.
DR PROSITE; PS01217; SUCCINYL-COA_LIG_3; 1.
KW Ligase; Tricarboxylic acid cycle; Complete proteome.
SQ SEQUENCE 382 AA; 42204 MW; 418ACBA41E64DA9F CRC64;
-----
Query Match 29.1%; Score 48; DB 1; Length 382;
Best Local Similarity 41.9%; Pred. No. 14;
Matches 13; Conservative 5; Mismatches 3; Indels 10; Gaps 2;

QY 4 KANS-----KFIGITELGGHEIKKVLV 25
| : : : | : | : |
Db 61 KANSPEAKEVAKKILGMT-IGHRVEKVLV 90
| : : : | : | : |

RESULT 15
LCAT_RAT
ID LCAT_RAT STANDARD; PRT; 440 AA.
AC P18424;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phosphatidylcholine-sterol acyltransferase precursor (EC 2.3.1.43)
DE (Lecithin-cholesterol acyltransferase) (Phospholipid-cholesterol
DE acyltransferase).
GN LCAT.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90384859; PubMed=2402469;
RA Meroni G., Margaretti N., Magnaghi P., Taramelli R.;
RT "Nucleotide sequence of the cDNA for lecithin-cholesterol acyl
RT transferase (LCAT) from the rat.";
RL Nucleic Acids Res. 18:5308-5308(1990).
CC -!- FUNCTION: CENTRAL ENZYME IN THE EXTRACELLULAR METABOLISM OF PLASMA
CC LIPOPROTEINS. AMONG OTHER SUBSTRATES IT ESTERIFIES THE FREE
CC CHOLESTEROL TRANSPORTED IN PLASMA LIPOPROTEINS.
CC -!- CATALYTIC ACTIVITY: Phosphatidylcholine + a sterol = a sterol
CC ester + 1-acylglycerophosphocholine.
CC -!- ENZYME REGULATION: APOLIPOPROTEIN A-I IS A POTENT ACTIVATOR FOR
CC THIS ENZYME.
CC -!- SIMILARITY: PARTIAL WITH LIPASES (PANCREATIC, GASTRIC, HEPATIC,
CC LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; X54096; CAA38030.1; -  
DR PIR; S11214; XRRTN.  
DR InterPro; IPR003386; LACT.  
DR InterPro; IPR00734; Lipase.  
DR InterPro; IPR00379; Ser\_estrs\_site.  
DR Pfam; PF02450; LACT; 1.  
DR PROSITE; PS00120; LIPASE\_SER; 1.  
KW Transferase; Acyltransferase; Lipid metabolism; Glycoprotein; Signal.  
FT SIGNAL 1 24  
FT CHAIN 25 440  
FT PHOSPHATIDYLCHOLINE-STEROL  
FT ACYLTRANSFERASE.  
FT ACT\_SITE 205 205  
FT DISULFID 74 98  
FT DISULFID 337 380  
FT CARBOHYD 44 44  
FT CARBOHYD 108 108  
FT CARBOHYD 296 296  
FT CARBOHYD 397 397  
FT CARBOHYD 408 408  
SQ SEQUENCE 440 AA; 49727 MW; 65E39212168A8885 CRC64;  
CHARGE RELAY SYSTEM (BY SIMILARITY).  
BY SIMILARITY.  
BY SIMILARITY.  
N-LINKED (GLCNAC. . .) (POTENTIAL).  
N-LINKED (GLCNAC. . .) (POTENTIAL).  
N-LINKED (GLCNAC. . .) (POTENTIAL).  
N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match 29.1%; Score 48; DB 1; Length 440;  
Best Local Similarity 50.0%; Pred. No. 16;  
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 14 ELGGHEIKKVLVPGCHGS 31  
||| :|||||:  
Db 41 ELSNHTRPVILVPGMG 58

Search completed: December 4, 2002, 12:53:27  
Job time : 5.35833 secs



```

Q9BIX2
ID Q9BIX2 PRELIMINARY; PRT; 170 AA.
AC Q9BIX2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 19, Last annotation update)
DE Allergen Def f II (Fragment).
OS Dermatophagoides farinae (House-dust mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcoptiformes; Astigmata; Analgoidea; Pyroglyphidae;
OC Dermatophagoides.
OX NCBI_TaxID=6954;
RN [1]
RP SEQUENCE FROM N.A.
RA Hao M., Xu J., Zhong N.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF346905; AAK30133.1; -
DR HSSP: Q00855; 1AHK.
DR InterPro: IPR003172; EL_DerP2_DerF2.
DR Pfam: PF02221; EL_DerP2_DerF2; 1.
FT NON_TER 1
SQ SEQUENCE 170 AA; 18781 MW; 0C2B58734C9D443A CRC64;

Query Match 43.6%; Score 72; DB 5; Length 170;
Best Local Similarity 50.0%; Pred. No. 0.0095;
Matches 14; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 4 KANSKEIGITELGGHRIKKVLVPGCHGS 31
| | | | | : | | | | | | | | | |
Db 38 KHNFLFLVYIHIANNEIKKVVMDGCHGS 65

RESULT 3
Q8WQK5
ID Q8WQK5 PRELIMINARY; PRT; 129 AA.
AC Q8WQK5;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Major Der f 2 isoform (Fragment).
OS Dermatophagoides farinae (House-dust mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcoptiformes; Astigmata; Analgoidea; Pyroglyphidae;
OC Dermatophagoides.
OX NCBI_TaxID=6954;
RN [1]
RP SEQUENCE FROM N.A.
RA Jin H.S., Oh S.H., Hong C.-S.;
RT "cDNA sequence encoding major Der f 2 isoform in Korea.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AV066008; AAL47677.1; -
DR InterPro: IPR003172; EL_DerP2_DerF2.
DR Pfam: PF02221; EL_DerP2_DerF2; 1.
FT NON_TER 1
SQ SEQUENCE 129 AA; 14035 MW; 832F72E25FE4F43F CRC64;

Query Match 40.6%; Score 67; DB 5; Length 129;
Best Local Similarity 50.0%; Pred. No. 0.039;
Matches 11; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 10 IGITELGGHRIKKVLVPGCHGS 31
| | | | | : | | | | | | | | | |
Db 3 VDVKDCANNEIKKVVMDGCHGS 24

RESULT 4
Q24262
ID Q24262 PRELIMINARY; PRT; 1333 AA.
AC Q24262;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE BLASTOPIA polyprotein.

```

```

GN BLASTOPIA\POLYPROTEIN.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OREGON R;
RX MEDLINE=94333069; PubMed=8055714;
RA Frommer G., Schuh R., Jackle H.;
RT "Localized expression of a novel microRNA-like element in the
RT blastoderm of Drosophila melanogaster is dependent on the anterior
RT morphogen bicoid.";
RL Chromosoma 103:82-89(1994).
DR EMBL: Z27119; CAA81643.1; -
DR Flybase; FBgn0043491; blastopia\polyprotein.
DR InterPro: IPR001995; Aspprotease_rtrv.
DR InterPro: IPR001969; Aspprotease_site.
DR InterPro: IPR001584; Rve.
DR InterPro: IPR000477; RVTse.
DR InterPro: IPR001878; Znf_CCHC.
DR Pfam: PF00665; rve; 1.
DR Pfam: PF00077; rvp; 1.
DR Pfam: PF00078; rvt; 1.
DR Pfam: PF00098; zf-CCHC; 2.
DR PRINTS; PR00939; C2HCZNFINGER.
DR SMART; SM00343; ZNF_C2HC; 2.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.
DR PROSITE; PS00175; ASP_RETROV; 1.
KW Polyprotein; RNA-directed DNA polymerase.
SQ SEQUENCE 1333 AA; 152451 MW; 2383E01108216E36 CRC64;

Query Match 32.7%; Score 54; DB 5; Length 1333;
Best Local Similarity 35.9%; Pred. No. 45;
Matches 14; Conservative 7; Mismatches 10; Indels 8; Gaps 1;

QY 1 QYIKANSKFGITEL-----GGHEIKKVLVPGCHGS 31
| | | | | | | | | | | | | | | |
Db 127 KYVOARSKMIGSAELFLESECVGYTELKELLIEFSGS 165

RESULT 5
Q9HRS0
ID Q9HRS0 PRELIMINARY; PRT; 230 AA.
AC Q9HRS0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Vng0573C.
GN Vng0573C.
OS Halobacterium sp. (strain NRC-1).
OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL; AE005007; AAG19088.1; -
DR HSSP; P50057; 2B31.
DR InterPro: IPR000923; BlueCu_1.
DR InterPro: IPR001235; Copper_bind; 1.
DR Pfam: PF00127; copper-bind; 1.

```

RESULT 7	
Q31585	
ID	Q31585 PRELIMINARY; PRT; 60 AA.
AC	A31585;
DT	01-NOV-1996 (TrEMBLrel. 01, Created)
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE	(DB02) MHC class II beta 1 (Fragment).
OS	Salmo salar (Atlantic salmon).
QC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OX	Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
NCBI_TaxID=8030;	
[1]	SEQUENCE FROM N.A.
RA	Grimholt U., Olsaker I., de Vries Lindstrom C., Lie O.;
RT	"A study of polymorphism in the MHC class II beta 1 and
RT	alpha 2 domain exons of Atlantic salmon (Salmo salar).";
RL	Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
DR	EMBL; L24953; AAA49597.1; -.
DR	InterPro; IPR000353; MHC_II_beta.
DR	Pfam; PF00969; MHC_II_beta; 1.
DR	ProDom; PD000328; MHC_II_beta; 1.

```

RESULT 9
Q95IS2
ID Q95IS2 PRELIMINARY; PRT; 85 AA.
AC Q95IS2;
DT 01-DEC-2001 (TtEMBLrel. 19, Created)
DT 01-DEC-2001 (TtEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TtEMBLrel. 20, Last annotation update)
DE MHC class II beta chain (Fragment).
OS Salmo salar (Atlantic salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX NCBI_TaxID=8030;
RN [1]
RP SEQUENCE FROM N.A.
RA Landry C., Bernatchez L.;
RT "Comparative analysis of population structure across environments and
RT geographic scales at Major Histocompatibility Complex and
RT microsatellite in Atlantic salmon (Salmo salar).";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF373699; AAK61882.1; -
DR InterPro; IPR000353; MHC II beta

```

```

DR Pfam: PF00969; MHC_II_beta; 1.
DR ProDom: PD000328; MHC_II_beta; 1.
KW Glycoprotein; MHC; MHC II; Transmembrane.
FT NON_TER 1
FT NON_TER 85
FT NON_TER 85
SQ SEQUENCE 85 AA; 9743 MW; 3214E01AD1B66AC5 CRC64;

Query Match 30.6%; Score 50.5; DB 7; Length 85;
Best Local Similarity 57.9%; Pred. No. 7.2;
Matches 11; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

Qy 1 QYIKANS---KFIGITELG 16
:|:|: || ||:| ||||
Db 33 EYIRFNSTVGKFGVGYTELG 51

RESULT 10
Q95HY1 PRELIMINARY; PRT; 85 AA.
AC Q95HY1;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE MHC class II B antigen (Fragment).
GN DBI.
OS Salmo salar (Atlantic salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX NCBI_TaxID=8030;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21383619; PubMed=11491536;
RA Langeforß A., Lohm J., von Schantz T.;
RT "Allelic polymorphism in MHC class II B in four populations of
RL Atlantic salmon (Salmo salar).";
DR EMBL; AF104370; AAL04002.1; -.
DR InterPro: IPR000353; MHC_II_beta.
DR Pfam: PF00969; MHC_II_beta; 1.
DR ProDom: PD000328; MHC_II_beta; 1.
KW Glycoprotein; MHC; MHC II; Transmembrane.
FT NON_TER 1
FT NON_TER 85
FT NON_TER 85
SQ SEQUENCE 85 AA; 9723 MW; 27C9F7931F1F01C4 CRC64;

Query Match 30.6%; Score 50.5; DB 7; Length 85;
Best Local Similarity 57.9%; Pred. No. 7.2;
Matches 11; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

Qy 1 QYIKANS---KFIGITELG 16
:|:|: || ||:| ||||
Db 33 EYIRFNSTVGKFGVGYTELG 51

RESULT 11
Q95HX4 PRELIMINARY; PRT; 86 AA.
AC Q95HX4;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE MHC class II B antigen (Fragment).
GN DBI.
OS Salmo salar (Atlantic salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX NCBI_TaxID=8030;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21383619; PubMed=11491536;
RA Langeforß A., Lohm J., von Schantz T.;

```

```

RT "Allelic polymorphism in MHC class II B in four populations of
RL Atlantic salmon (Salmo salar).";
RL Immunogenetics 53:329-336(2001).
DR EMBL; AF104377; AAL04009.1; -.
DR InterPro: IPR000353; MHC_II_beta.
DR Pfam: PF00969; MHC_II_beta; 1.
DR ProDom: PD000328; MHC_II_beta; 1.
KW Glycoprotein; MHC; MHC II; Transmembrane.
FT NON_TER 1
FT NON_TER 86
FT NON_TER 86
SQ SEQUENCE 86 AA; 9912 MW; E5097729F681F149 CRC64;

Query Match 30.6%; Score 50.5; DB 7; Length 86;
Best Local Similarity 57.9%; Pred. No. 7.3;
Matches 11; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

Qy 1 QYIKANS---KFIGITELG 16
:|:|: || ||:| ||||
Db 33 EYIRFNSTVGKFGVGYTELG 51

RESULT 12
Q31590 PRELIMINARY; PRT; 244 AA.
AC Q31590;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MHC class II.
GN MHC-SASA CLASS II B.
OS Salmo salar (Atlantic salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX NCBI_TaxID=8030;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=LEUKOCYTE;
RX MEDLINE=93170890; PubMed=8436418;
RA Hordvik I., Grimholt U., Fosse V.M., Lie Y., Endresen C.;
RT "Cloning and sequence analysis of cDNAs encoding the MHC class II a-
chain in Atlantic salmon, Salmo salar.";
RL Immunogenetics 37:437-441(1993).
DR EMBL; X70166; CAA49725.1; -.
DR HSSP; P01888; IBMG.
DR InterPro: IPR003597; Ig_cl.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR000353; MHC_II_beta.
DR Pfam: PF00047; Ig; 1.
DR Pfam: PF00969; MHC_II_beta; 1.
DR ProDom: PD000328; MHC_II_beta; 1.
DR SMART; SM00407; IGL1; 1.
KW Glycoprotein; MHC II; Transmembrane.
SQ SEQUENCE 244 AA; 27449 MW; 496CB9EA9D73765C CRC64;

Query Match 30.6%; Score 50.5; DB 7; Length 244;
Best Local Similarity 57.9%; Pred. No. 23;
Matches 11; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

Qy 1 QYIKANS---KFIGITELG 16
:|:|: || ||:| ||||
Db 51 EYIRFNSTVGKFGVGYTELG 69

RESULT 13
Q9PAB9 PRELIMINARY; PRT; 433 AA.
AC Q9PAB9;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Phosphodiesterase-nucleotide pyrophosphatase precursor.
GN XF2599.

```



1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	8												

Db 77 KATGRFVG--EAGPHDLKRDMEPSIEG 101

Search completed: December 4, 2002, 12:52:55  
Job time : 14.0125 secs

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 12:49:05 ; Search time 15,3708 Seconds  
(without alignments)  
268,741 Million cell updates/sec

Title: US-09-362-731A-1  
Perfect score: 165  
Sequence: 1 QVIRANSKFIGITELGGHEIKKVLPGCHGS 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

```

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	165	100.0	31	21	Tetanus toxoid T c
2	102.5	62.1	137	21	Tetanus toxoid T c
3	98	59.4	32	21	Tetanus toxoid T c
4	92	55.8	47	15	LHRH-containing im
5	86	52.1	29	16	IgE CH4 region con
6	86	52.1	37	15	Universal immunost
7	86	52.1	37	15	Universal immunost
8	85	51.5	129	23	House dust mite al
9	85	51.5	129	23	House dust mite al
10	84	50.9	25	16	Dermatophagoides p

11	84	50.9	25	19	AAW71901	Dermatophagoides p
12	84	50.9	25	20	AAV50520	Dermatophagoides s
13	84	50.9	25	22	AAU19123	T-cell epitope con
14	84	50.9	26	14	AAR36424	Der p II derived p
15	84	50.9	26	15	AAR51772	Der p II derived p
16	84	50.9	26	19	AAW72303	Dermatophagoides D
17	84	50.9	26	19	AAW71947	Dermatophagoides D
18	84	50.9	26	20	AAV50401	Dermatophagoides s
19	84	50.9	26	20	AAV50521	Dermatophagoides s
20	84	50.9	26	22	AAU19004	T-cell epitope con
21	84	50.9	26	22	AAU19124	T-cell epitope con
22	84	50.9	27	19	AAW72304	Dermatophagoides D
23	84	50.9	27	20	AAV50522	Dermatophagoides s
24	84	50.9	27	22	AAU19125	T-cell epitope con
25	84	50.9	129	15	AAAR49921	Protein allergen o
26	84	50.9	129	15	AAAR49922	Protein allergen o
27	84	50.9	129	22	AAU80061	Modified tick alle
28	84	50.9	129	22	AAAG78996	House dust mite al
29	84	50.9	129	22	AAU07751	House dust mite al
30	84	50.9	129	23	ABG66973	House dust mite al
31	84	50.9	129	23	ABG66974	House dust mite al
32	84	50.9	129	23	ABG66975	House dust mite al
33	84	50.9	129	23	ABG66976	House dust mite al
34	84	50.9	129	23	ABG66991	House dust mite al
35	84	50.9	129	23	ABG66992	House dust mite al
36	84	50.9	129	23	ABG66993	House dust mite al
37	84	50.9	129	23	ABG66994	House dust mite al
38	84	50.9	129	23	ABG66995	House dust mite al
39	84	50.9	129	23	ABG66996	House dust mite al
40	84	50.9	129	23	ABG66997	House dust mite al
41	84	50.9	129	23	ABG66999	House dust mite al
42	84	50.9	129	23	ABG67002	House dust mite al
43	84	50.9	129	23	ABG67003	House dust mite al
44	84	50.9	129	23	ABG67004	House dust mite al
45	84	50.9	129	23	ABG67005	House dust mite al

#### ALIGNMENTS

RESULT 1  
AAV82632  
ID AAV82632 standard; peptide; 31 AA.  
XX  
AC AAV82632;  
XX  
DT 07-AUG-2000 (first entry)  
XX  
DE Tetanus toxoid T cell epitope and Der pII B cell epitope peptide.  
XX  
KW T cell epitope; B cell epitope; allergy; allergen; antigenic;  
KW antiallergic; antiasthmatic; antiinflammatory; dermatological;  
KW immunosuppressive; vaccine; rhinitis; sinusitis; bronchial asthma;  
KW atopic dermatitis; acute urticaria; chronic urticaria;  
KW gastro-intestinal syndrome; food allergen; oro-pharyngeal syndrome;  
KW anaphylactic reaction; drug hypersensitivity; allergic reaction.  
XX  
OS Dermatophagoides pteronyssinus.  
OS Clostridium tetani.  
OS Synthetic.  
XX  
XX WO200006694-A2.  
XX  
PD 10-FEB-2000.  
XX  
PF 20-JUL-1999; 99WO-BE00092.  
XX  
PR 30-JUL-1998; 98EP-0870167.  
XX  
PA (UNIO ) UCB SA.  
XX  
PI Saint-Remy J, Jacquemin M;  
XX

```

DR  WPI; 2000-422470/36.
XX
PT  New compound for prevention and treatment of allergies comprises at
PT  least one allergen antigenic determinant recognized by a B cell and at
PT  least one antigenic determinant which does not trigger T cell
PT  activation -
XX
PS  Claim 8; Page 35; 50pp; English.
XX
CC  The present invention describes a compound (I) for the prevention and/or
CC  treatment of allergy. The compound comprises at least one allergen
CC  antigenic determinant (i) recognised by a B cell or an antibody secreted
CC  by a B cell of a non-atopic individual and at least one antigenic
CC  determinant (ii) different from the allergen that triggers T cell
CC  activation. (I) has antiallergic, antiasthmatic, antiinflammatory,
CC  dermatological and immunosuppressive activities, and can be used in a
CC  vaccine. (I) may be used in a pharmaceutical or cosmetic medicament to
CC  treat and/or prevent allergies or a disease of allergic origin,
CC  especially hypersensitivities. These include rhinitis, sinusitis,
CC  bronchial asthma, atopic dermatitis, some forms of acute and chronic
CC  urticaria, gastro-intestinal syndromes associated with the ingestion of
CC  food allergens, oro-pharyngeal syndrome, anaphylactic reactions
CC  associated with drug hypersensitivities and/or a mixture of these. The
CC  use of (I) in the treatment of allergic conditions avoids the need for
CC  drug treatment, which often causes undesirable side-effects. Also, prior
CC  art drug therapies alleviate symptoms, but do not influence their
CC  causes, however (I) actually combats the cause of an allergic reaction.
CC  The present sequence represents a specifically claimed compound peptide
CC  sequence from the present invention.
XX
SQ  Sequence 31 AA;
    Query Match      100.0%; Score 165; DB 21; Length 31;
    Best Local Similarity 100.0%; Pred. No. 5.2e-19;
    Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 QYIKANSKFGITELGGHEIKKVLVPGCHGS 31
    |||||
DB  1 QYIKANSKFGITELGGHEIKKVLVPGCHGS 31

RESULT 2
AAV82634
ID  AAY82634 standard; peptide; 137 AA.
XX
AC  AAY82634;
XX
DT  07-AUG-2000 (first entry)
XX
DE  Tetanus toxoid T cell epitopes and Der pII B cell epitopes peptide.
XX
KW  T cell epitope; B cell epitope; allergy; allergen; antigenic;
KW  antiallergic; antiasthmatic; antiinflammatory; dermatological;
KW  immunosuppressive; vaccine; rhinitis; sinusitis; bronchial asthma;
KW  atopic dermatitis; acute urticaria; chronic urticaria;
KW  gastro-intestinal syndrome; food allergen; oro-pharyngeal syndrome;
KW  anaphylactic reaction; drug hypersensitivity; allergic reaction.
XX
OS  Dermatophagoides pteronyssinus.
OS  Clostridium tetani.
OS  Synthetic.
XX
PN  WO200006694-A2.
XX
PD  10-FEB-2000.
XX
PF  20-JUL-1999; 99WO-BE00092.
XX
PR  30-JUL-1998; 98EP-0870167.
XX
PA  (UNIO ) UCB SA.
XX
PI  Saint-Remy J, Jacquemin M;

```

---

```

XX
DR  WPI; 2000-422470/36.
XX
PT  New compound for prevention and treatment of allergies comprises at
PT  least one allergen antigenic determinant recognized by a B cell and at
PT  least one antigenic determinant which does not trigger T cell
PT  activation -
XX
PS  Claim 8; Page 35; 50pp; English.
XX
CC  The present invention describes a compound (I) for the prevention and/or
CC  treatment of allergy. The compound comprises at least one allergen
CC  antigenic determinant (i) recognised by a B cell or an antibody secreted
CC  by a B cell of a non-atopic individual and at least one antigenic
CC  determinant (ii) different from the allergen that triggers T cell
CC  activation. (I) has antiallergic, antiasthmatic, antiinflammatory,
CC  dermatological and immunosuppressive activities, and can be used in a
CC  vaccine. (I) may be used in a pharmaceutical or cosmetic medicament to
CC  treat and/or prevent allergies or a disease of allergic origin,
CC  especially hypersensitivities. These include rhinitis, sinusitis,
CC  bronchial asthma, atopic dermatitis, some forms of acute and chronic
CC  urticaria, gastro-intestinal syndromes associated with the ingestion of
CC  food allergens, oro-pharyngeal syndrome, anaphylactic reactions
CC  associated with drug hypersensitivities and/or a mixture of these. The
CC  use of (I) in the treatment of allergic conditions avoids the need for
CC  drug treatment, which often causes undesirable side-effects. Also, prior
CC  art drug therapies alleviate symptoms, but do not influence their
CC  causes, however (I) actually combats the cause of an allergic reaction.
CC  The present sequence represents a specifically claimed compound peptide
CC  sequence from the present invention.
XX
SQ  Sequence 137 AA;
    Query Match      62.1%; Score 102.5; DB 21; Length 137;
    Best Local Similarity 60.5%; Pred. No. 2.6e-08;
    Matches 23; Conservative 1; Mismatches 7; Indels 7; Gaps 1;

QY  1 QYIKANSKFGITELGGHEIKK-----VLVPGCHGS 31
    |||||
DB  2 QYIKANSKFGITELGGQYIKANSKFGITELSSCHGS 39

RESULT 3
AAV82636
ID  AAY82636 standard; peptide; 32 AA.
XX
AC  AAY82636;
XX
DT  07-AUG-2000 (first entry)
XX
DE  Tetanus toxoid T cell epitope and Der pII B cell epitope peptide.
XX
KW  T cell epitope; B cell epitope; allergy; allergen; antigenic;
KW  antiallergic; antiasthmatic; antiinflammatory; dermatological;
KW  immunosuppressive; vaccine; rhinitis; sinusitis; bronchial asthma;
KW  atopic dermatitis; acute urticaria; chronic urticaria;
KW  gastro-intestinal syndrome; food allergen; oro-pharyngeal syndrome;
KW  anaphylactic reaction; drug hypersensitivity; allergic reaction.
XX
OS  Dermatophagoides pteronyssinus.
OS  Clostridium tetani.
OS  Synthetic.
XX
PN  WO200006694-A2.
XX
PD  10-FEB-2000.
XX
PF  20-JUL-1999; 99WO-BE00092.
XX
PR  30-JUL-1998; 98EP-0870167.
XX
PA  (UNIO ) UCB SA.
XX

```

PI Saint-Remy J, Jacquemin M;  
 DR WPI; 2000-422470/36.  
 XX  
 XX  
 PT New compound for prevention and treatment of allergies comprises at  
 PT least one allergen antigenic determinant recognized by a B cell and at  
 PT least one antigenic determinant which does not trigger T cell  
 PT activation -  
 XX  
 XX  
 PS Claim 8; Page 35; 50pp; English.  
 XX  
 XX The present invention describes a compound (I) for the prevention and/or  
 CC treatment of allergy. The compound comprises at least one allergen  
 CC antigenic determinant (i) recognised by a B cell or an antibody secreted  
 CC by a B cell of a non-atopic individual and at least one antigenic  
 CC determinant (ii) different from the allergen that triggers T cell  
 CC activation. (I) has antiallergic, antiasthmatic, antiinflammatory,  
 CC dermatological and immunosuppressive activities, and can be used in a  
 CC vaccine. (I) may be used in a pharmaceutical or cosmetic medicament to  
 CC treat and/or prevent allergies or a disease of allergic origin,  
 CC especially hypersensitivities. These include rhinitis, sinusitis,  
 CC bronchial asthma, atopic dermatitis, some forms of acute and chronic  
 CC urticaria, gastro-intestinal syndromes associated with the ingestion of  
 CC food allergens, oro-pharyngeal syndrome, anaphylactic reactions  
 CC associated with drug hypersensitivities and/or a mixture of these. The  
 CC use of (I) in the treatment of allergic conditions avoids the need for  
 CC drug treatment, which often causes undesirable side-effects. Also, prior  
 CC art drug therapies alleviate symptoms, but do not influence their  
 CC causes, however (I) actually combats the cause of an allergic reaction.  
 CC The present sequence represents a specifically claimed compound peptide  
 CC sequence from the present invention.  
 XX  
 XX Sequence 32 AA;  
 SQ  
 Query Match 59.4%; Score 98; DB 21; Length 32;  
 Best Local Similarity 67.7%; Pred. No. 2.7e-08;  
 Matches 21; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
 QY 1 QYIKANSKFIGITELGGHEIKKVLVPGCHGS 31  
 DB 1 QYIKANSKFIGITELG-----GCHGS 21  
 RESULT 4  
 AAR62723  
 ID AAR62723 standard; peptide: 47 AA.  
 XX  
 XX AAR62723;  
 AC  
 XX  
 XX 17-SEP-1995 (first entry)  
 DE LHRH-containing immunogenic peptide.  
 XX  
 XX Helper T cell epitope; universal immune stimulator; invasin; hapten;  
 KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;  
 KW androgen-dependent carcinoma; antitumour; infertility;  
 KW tetanus toxin.  
 XX  
 XX Synthetic.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 1..16  
 FT /note= "invasin domain"  
 FT 19..35  
 FT /note= "tetanus toxin helper T cell epitope"  
 FT 38..47  
 FT /note= "LHRH hapten"  
 XX  
 XX WO9425060-A.  
 PN  
 XX 10-NOV-1994.  
 PD  
 XX  
 XX 28-APR-1994; 94WO-US04832.  
 PF

XX 27-APR-1993; 93US-0057166.  
 PR 14-APR-1994; 94US-0229275.  
 XX  
 XX (LADD/) LADD A E.  
 PA (WANG/) WANG C Y.  
 PA (ZAMB/) ZAMB T.  
 XX  
 XX Ladd AE, Wang CY, Zamb T;  
 PI  
 XX WPI; 1994-357910/44.  
 DR  
 XX Immunogenic luteinising hormone releasing hormone peptide(s) -  
 PT that suppress LHRH activity in males and females  
 PT  
 XX Claim 8; Page 88; 213pp; English.  
 PS  
 XX Synthetic immunogenic peptides are provided in which a universal immune  
 CC stimulator is linked to a peptide or protein hapten containing B cell  
 CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes  
 CC potent immune responses to the coupled peptide or protein. The  
 CC stimulator consists of (A) a promiscuous helper T cell epitope (Th)  
 CC which elicits an immune response to the coupled peptide in members of  
 CC a heterogeneous population expressing diverse HLA phenotypes, and (B)  
 CC an adjuvant peptide sequence from the invasin protein of Yersinia.  
 CC Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the  
 CC invasin and Th domains and between the immune stimulator and hapten  
 CC components. When the hapten is LHRH, then optionally the invasin domain  
 CC can be omitted from the immune stimulator component.  
 CC The present sequence represents an LHRH-containing immunogenic peptide  
 CC as above which can be used as a potent vaccine for treating e.g.  
 CC prostatic hyperplasia, androgen-dependent carcinoma, prostatic  
 CC carcinoma, testicular carcinoma, endometriosis, benign uterine tumours,  
 CC recurrent functional ovarian cysts, (severe) premenstrual syndrome or  
 CC oestrogen-dependent breast cancer, or for induction of infertility.  
 XX  
 XX Sequence 47 AA;  
 SQ  
 Query Match 55.8%; Score 92; DB 15; Length 47;  
 Best Local Similarity 74.1%; Pred. No. 3.8e-07;  
 Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 1 QYIKANSKFIGITELGGHEIKKVLVPG 27  
 DB 21 QYIKANSKFIGITELGGHEHWSYGLRPG 47  
 RESULT 5  
 AAR83561  
 ID AAR83561 standard; peptide: 29 AA.  
 XX  
 XX AAR83561;  
 AC  
 XX  
 XX 13-JUN-1996 (first entry)  
 DE IgE CH4 region contg. peptide immunogen for treating allergies.  
 XX  
 XX IgE; CH4; immunoglobulin; epsilon; immunogen; helper T cell; epitope;  
 KW vaccine; allergy; antibody; constant heavy chain.  
 KW  
 XX Synthetic.  
 OS  
 XX  
 XX WO9526365-A1.  
 PN  
 XX  
 PD 05-OCT-1995.  
 XX  
 XX 24-MAR-1995; 95WO-US03741.  
 PF  
 XX 25-OCT-1994; 94US-0328912.  
 PR  
 XX 28-MAR-1994; 94US-0218461.  
 XX  
 XX (UNBI-) UNITED BIOMEDICAL INC.  
 PA  
 XX

PI Wang CY;  
 XX WPI; 1995-351297/45.  
 XX Synthetic peptide-based immunogen contg. IgE CH4 peptide and helper  
 PT T cell epitope - useful for eliciting antibody prodn. for allergy  
 PT treatment  
 XX  
 PS Claim 5; Page 68-69; 87pp; English.  
 XX  
 CC AAR82592-R82600 and AAR83560-R83581 are peptide immunogens that are  
 CC useful in vaccines for treating allergic reactions. In the immunogens,  
 CC an iGE CH4 peptide is attached C-terminally to a series of amino acids  
 CC including a helper T cell epitope. The immunogen may also opt. contain  
 CC a fatty acid or fatty acid derivative, an invasin domain or alpha-NH2.  
 CC The immunogen produces high titres of antibodies to the effector site  
 CC in human iGE heavy chain (the CH4 domain peptide) which inhibit mast  
 CC cell activation and reduce allergen-induced iGE prodn. The immunogens  
 CC may be used in either a radially branching multimeric form or a  
 CC linearly arranged monomeric form.  
 XX  
 SQ Sequence 29 AA;  
 Query Match 52.1%; Score 86; DB 16; Length 29;  
 Best Local Similarity 100.0%; Pred. No. 2e-06;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QYIKANSKFIGITELGG 17  
 Db 3 QYIKANSKFIGITELGG 19  
 RESULT 6  
 AAR65389  
 ID AAR65389 standard; peptide; 37 AA.  
 AC AAR65389;  
 XX  
 XX 21-SEP-1995 (first entry)  
 DT  
 XX Universal immunostimulator having GG spacers.  
 DE  
 XX Helper T cell epitope; universal immune stimulator; invasin; haptens;  
 KW tetanus toxin.  
 KW  
 XX Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 FH Domain 3..19  
 FT Domain /note= "tetanus toxin helper T cell epitope"  
 FT Domain 22..37  
 FT Domain /note= "invasin domain"  
 FT  
 XX WO9425060-A.  
 PN 10-NOV-1994.  
 PD  
 XX 28-APR-1994; 94WO-US04832.  
 XX  
 XX 27-APR-1993; 93US-0057166.  
 PR 14-APR-1994; 94US-0229275.  
 XX  
 XX (LADD/) LADD A E.  
 PA (WANG/) WANG C Y.  
 PA (ZAMB/) ZAMB T.  
 XX  
 PI Ladd AE, Wang CY, Zamb T;  
 XX WPI; 1994-357910/44.  
 XX  
 PT Immunogenic luteinising hormone releasing hormone peptide(s) -  
 PT that suppress LHRH activity in males and females  
 XX

PS Disclosure; Page 95; 213pp; English.  
 XX  
 CC Synthetic immunogenic peptides are provided in which a universal immune  
 CC stimulator is linked to a peptide or protein haptens containing B cell  
 CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes  
 CC potent immune responses to the coupled peptide or protein. The  
 CC stimulator consists of (A) a promiscuous helper T cell epitope (Th)  
 CC which elicits an immune response to the coupled peptide in members of  
 CC a heterogeneous population expressing diverse HLA phenotypes, and (B)  
 CC an adjuvant peptide sequence from the invasin protein of Yersinia.  
 CC Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the  
 CC invasin and Th domains and between the immune stimulator and haptens  
 CC components. When the haptens is LHRH, then optionally the invasin domain  
 CC can be omitted from the immune stimulator component.  
 CC The present sequence is an example of a -GG-Th-GG-invasin immune  
 CC stimulator to which a haptens can be bonded.  
 XX  
 SQ Sequence 37 AA;  
 Query Match 52.1%; Score 86; DB 15; Length 37;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-06;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QYIKANSKFIGITELGG 17  
 Db 5 QYIKANSKFIGITELGG 21  
 RESULT 7  
 AAR65383  
 ID AAR65383 standard; peptide; 37 AA.  
 AC AAR65383;  
 XX  
 XX 21-SEP-1995 (first entry)  
 DT  
 XX Universal immunostimulator having GG spacers.  
 DE  
 XX Helper T cell epitope; universal immune stimulator; invasin; haptens;  
 KW tetanus toxin.  
 KW  
 XX Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 FH Domain 1..16  
 FT Domain /note= "invasin domain"  
 FT Domain 19..35  
 FT Domain /note= "tetanus toxin helper T cell epitope"  
 FT  
 XX WO9425060-A.  
 PN 10-NOV-1994.  
 PD  
 XX 28-APR-1994; 94WO-US04832.  
 XX  
 XX 27-APR-1993; 93US-0057166.  
 PR 14-APR-1994; 94US-0229275.  
 XX  
 XX (LADD/) LADD A E.  
 PA (WANG/) WANG C Y.  
 PA (ZAMB/) ZAMB T.  
 XX  
 PI Ladd AE, Wang CY, Zamb T;  
 XX WPI; 1994-357910/44.  
 XX  
 PT Immunogenic luteinising hormone releasing hormone peptide(s) -  
 PT that suppress LHRH activity in males and females  
 XX  
 PS Disclosure; Page 95; 213pp; English.  
 XX  
 CC Synthetic immunogenic peptides are provided in which a universal immune  
 CC stimulator is linked to a peptide or protein haptens containing B cell

CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes  
 CC potential immune responses to the coupled peptide or protein. The  
 CC stimulator consists of (A) a promiscuous helper T cell epitope (Th)  
 CC which elicits an immune response to the coupled peptide in members of  
 CC a heterogeneous population expressing diverse HLA phenotypes, and (B)  
 CC an adjuvant peptide sequence from the invasive protein of Yersinia.  
 CC Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the  
 CC invasive and Th domains and between the immune stimulator and hapten  
 CC components. When the hapten is LHRH, then optionally the invasive domain  
 CC can be omitted from the immune stimulator component.  
 CC The present sequence is an example of an invasive-GG-Th-GG- immune  
 CC stimulator to which a hapten can be bonded.

XX  
 SQ Sequence 37 AA;

Query Match 52.1%; Score 86; DB 15; Length 37;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-06;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITELGG 17  
 |||||  
 Db 21 QYIKANSKFIGITELGG 37

RESULT 8  
 ID ABG66971 standard; Protein; 129 AA.

XX AC ABG66971;

DT 24-SEP-2002 (first entry)

XX House dust mite allergen Der p 2 mutant K6A.

XX Immunoglobulin E; IgE; allergen; allergy; mite; hay fever;  
 KW rhinoconjunctivitis; rhinitis; asthma; systemic anaphylaxis; mutant;  
 KW vaccine; antiallergic; B cell epitope.

XX Dermatophagoides pteronyssinus.  
 OS Synthetic.

XX WO200240676-A2.

XX 23-MAY-2002.

XX 16-NOV-2001; 2001WO-DK00764.

XX 16-NOV-2000; 2000DK-0001718.

PR 16-NOV-2000; 2000US-249361P.

PR 14-JUN-2001; 2001US-298170P.

XX (ALKA-) ALK-ABELLO AS.

XX Holm J, Ipsen H, Nedergaard Larsen J, Spangfort MD;

XX WPI; 2002-508328/54.

XX New recombinant mutant allergen, useful for preventing and/or treating  
 PT allergy, comprises multiple mutations and reduced immunoglobulin E  
 PT binding affinity -

XX Example 3; Page -: 210pp; English.

XX The invention relates to a recombinant allergen (I) which is a mutant of  
 CC a naturally occurring allergen, where the mutant allergen has at least  
 CC four primary mutations, which each reduce the specific immunoglobulin E  
 CC (IgE) binding capability of the mutated allergen as compared to the IgE  
 CC binding capability of the naturally occurring allergen, where each  
 CC primary mutation is a substitution of one surface-exposed amino acid  
 CC residue with another residue, which does not occur in the same position  
 CC in the amino acid sequence of any known homologous protein within the  
 CC taxonomic species from which the naturally occurring allergen  
 CC originates, and each primary mutation is spaced from each other primary

CC mutation by at least 15 Angstrom , and the primary mutations are placed  
 CC in such a manner that at least one circular surface region with a area  
 CC of 800 Angstrom ^2 comprises no mutation. Also included are a composition  
 CC comprising two or more of the recombinant allergens, where the variant  
 CC allergen is defined by having at least one primary mutation, which is  
 CC absent in at least one of the other variants, and for each variant no  
 CC secondary mutation is present within a radius of 15 Angstrom from each  
 CC absent primary mutation; a DNA sequence encoding the recombinant allergen  
 CC or its derivative, partial sequence or degenerated sequence, or a  
 CC sequence which hybridises to it under stringent conditions, where the  
 CC derivative, partial sequence, degenerated sequence or hybridising  
 CC sequence encodes a peptide having at least one B cell epitope; an  
 CC expression vector comprising the DNA and a host cell comprising the  
 CC vector. The recombinant allergen is useful as a pharmaceutical, for  
 CC preparing a pharmaceutical for preventing and/or treating allergy, or in  
 CC a diagnostic assay for assessing relevance, safety or outcome of therapy  
 CC of a subject, where an IgE containing sample of the subject is mixed  
 CC with the recombinant allergen and assessed for the level of reactivity  
 CC between the IgE in the sample and the recombinant allergen. The  
 CC recombinant allergen or compositions are useful for generating an  
 CC immune response in a subject, for vaccination or treatment of a subject  
 CC or for the treatment, prevention or alleviation of allergic reactions  
 CC in a subject e.g. hay fever, rhinoconjunctivitis, rhinitis, asthma or  
 CC systemic anaphylaxis. The present sequence represents a  
 CC recombinant allergen of the invention.

CC Note: The present sequence was not shown in the specification but  
 CC was created by the indexer using information in the specification and  
 CC the corresponding wild-type sequence.

SQ Sequence 129 AA;

Query Match 51.5%; Score 85; DB 23; Length 129;  
 Best Local Similarity 63.6%; Pred. No. 1.5e-05;  
 Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31

Db 3 VDVADCANHEIKKVLVPGCHGS 24

RESULT 9

ABG66998

ID ABG66998 standard; Protein; 129 AA.

XX AC ABG66998;

DT 24-SEP-2002 (first entry)

XX House dust mite allergen Der p 2 ALK-114 mutant K6A.

XX Immunoglobulin E; IgE; allergen; allergy; mite; hay fever;  
 KW rhinoconjunctivitis; rhinitis; asthma; systemic anaphylaxis; mutant;  
 KW vaccine; antiallergic; B cell epitope.

XX Dermatophagoides pteronyssinus.  
 OS Synthetic.

XX WO200240676-A2.

XX 23-MAY-2002.

XX 16-NOV-2001; 2001WO-DK00764.

XX 16-NOV-2000; 2000DK-0001718.

PR 16-NOV-2000; 2000US-249361P.

PR 14-JUN-2001; 2001US-298170P.

XX (ALKA-) ALK-ABELLO AS.

XX Holm J, Ipsen H, Nedergaard Larsen J, Spangfort MD;

XX WPI; 2002-508328/54.

XX





XX The present invention describes peptides for treating sensitivity to  
 CC house dust mite allergens from the genus Dermatophagoides. Peptides  
 CC within the scope of the invention comprise at least one T cell epitope,  
 CC or preferably at least two T cell epitopes of a protein allergen  
 CC selected from the allergens Der p I, Der p II, Der f I, or Der f II.  
 CC The invention also describes modified peptides having similar or  
 CC enhanced therapeutic properties as the corresponding, naturally  
 CC occurring allergen, but having reduced side effects. AAW71897 to  
 CC AAW71907 represent specifically claimed peptides.

XX SQ Sequence 25 AA;

Query Match 50.9%; Score 84; DB 19; Length 25;  
 Best Local Similarity 63.6%; Pred. No. 3.5e-06;  
 Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31  
 : : : |||||  
 Db 3 VDVKDCANHEIKKVLVPGCHGS 24

RESULT 12  
 AAY50520  
 ID AAY50520 standard; Peptide; 25 AA.  
 XX AC AAY50520;  
 XX DT 25-JAN-2000 (first entry)  
 XX DE Dermatophagoides sp major protein allergen DP11-20.9.  
 XX KW Allergen; house dust mite; detection; sensitivity; T cell epitope;  
 KW screening; allergic disorder; asthma; rhinitis; ectopic dermatitis;  
 KW Der f I; Der p I; Der p II; Der f II.  
 XX OS Dermatophagoides sp.  
 XX PN US5968526-A.  
 XX PR 19-OCT-1999.  
 XX PD 07-JUN-1995; 95US-0478572.  
 XX PF 19-MAY-1995; 95US-0445307.  
 XX PR 14-APR-1994; 94US-0227772.  
 XX PR 12-APR-1995; 95WO-US04481.  
 XX PA (IMMU-) IMMULOGIC PHARM CORP.  
 XX PI Garman RD, Greenstein JL, Rogers BL, Franzen HM, Shaked Z, Chen X;  
 PI Evans S, Kuo M;  
 DR WPI; 1999-590385/50.  
 XX Screening individuals for allergic reactions to T cell epitopes of  
 PT major allergens from house dust mites -  
 XX Claim 7; Column 165-166; 158pp; English.

XX This invention describes a novel method (I) for detecting whether an  
 CC individual is sensitive to Dermatophagoides (house dust mites). The  
 CC method involves detecting sensitivity to house dust mites in patients,  
 CC comprising combining a blood sample from the individual with 1 or more  
 CC isolated T cell epitopes of the protein allergens I and II ((DP I) and  
 CC (DP II)) from Dermatophagoides (house dust mites). 32 T cell epitopes  
 CC with varying, defined amino acids sequences (given in the specification)  
 CC may be used in (I). The sample and allergens are combined under  
 CC conditions appropriate for the binding of blood components with the  
 CC polypeptides. The extent of binding is then indicative of the  
 CC sensitivity of the patient to house dust mites. (I) may be used to screen  
 CC individuals for sensitivity to Dermatophagoides (house dust mites). The  
 CC house dust mite is a major cause of a variety of allergic disorders such

CC as asthma, rhinitis and ectopic dermatitis. AAY50360-Y50542 and  
 CC AAY50546-Y50555 represent house dust mite allergen peptide fragments  
 CC derived from Der p I, Der f II, Der f I and Der f II.

XX SQ Sequence 25 AA;

Query Match 50.9%; Score 84; DB 20; Length 25;  
 Best Local Similarity 63.6%; Pred. No. 3.5e-06;  
 Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31  
 : : : |||||  
 Db 3 VDVKDCANHEIKKVLVPGCHGS 24

RESULT 13  
 AAU19123  
 ID AAU19123 standard; Peptide; 25 AA.  
 XX AC AAU19123;  
 XX DT 04-DEC-2001 (first entry)  
 XX DE T-cell epitope containing peptide DP11-20.9.  
 XX KW House dust mite; allergenic peptide; Der p I; Der p II; Der f I;  
 KW Der f II; antiallergenic; immunostimulant; house dust mite allergy;  
 KW T-cell epitope.  
 XX OS Dermatophagoides pteronyssinus.  
 XX PN US6268491-B1.  
 XX PD 31-JUL-2001.  
 XX PF 07-JUN-1995; 95US-0484296.  
 XX PR 19-MAY-1995; 95US-0445307.  
 XX PR 16-OCT-1991; 91US-0777859.  
 XX PR 08-MAY-1992; 92US-0881396.  
 XX PR 14-APR-1993; 93WO-US03471.  
 XX PR 14-APR-1994; 94US-0227772.  
 XX PA (IMMU-) IMMULOGIC PHARM CORP.  
 XX PI Garman RD, Greenstein JL, Kuo M, Rogers BL, Franzen HM, Chen X;  
 PI Evans S, Shaked Z;  
 DR WPI; 2001-549074/61.  
 XX Peptides comprising T cell groups of the major allergens from  
 PT Dermatophagoides (house dust mites), useful for treating house dust  
 PT mite allergy in humans, and for diagnosing sensitivity to house dust  
 PT mite protein allergens -  
 XX Claim 2; Figure 30; 158pp; English.

XX The invention relates to an isolated peptide of the major protein  
 CC allergens of the genus Dermatophagoides, which comprises at least one T  
 CC cell group of a protein allergen from Der p (DP) I, DP II, Der f (DF) I  
 CC or DF II. The isolated peptide comprises at least two regions,  
 CC each region comprising at least one T cell group of a protein allergen  
 CC of the genus Dermatophagoides. The regions are derived from the same or  
 CC different protein allergens of the genus Dermatophagoides. The peptides  
 CC are useful for treating house dust mite allergy in humans. The peptides  
 CC are also useful for detecting or diagnosing sensitivity to house dust  
 CC mite protein allergens. The present peptides have similar or enhanced  
 CC therapeutic properties as the naturally-occurring allergen, but have  
 CC reduced side effects, and increased solubility and stability. The  
 CC present sequence represents an allergenic T-cell epitope containing  
 CC peptide derived from the Dermatophagoides allergenic proteins.

XX SQ Sequence 25 AA;

```

Query Match          50.9%; Score 84; DB 22; Length 25;
Best Local Similarity 63.6%; Pred. No. 3.5e-06;
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31
Db 3 VDVKDCANHEIKKVLVPGCHGS 24

RESULT 14
AAR36424
ID AAR36424 standard; peptide; 26 AA.
XX
AC AAR36424;
XX
DT 12-AUG-1993 (first entry)
XX
DE DPTI-20(1-26) a Dermatophagoides protein allergen.
XX
KW T cell epitope; house dust mite; allergy; soluble; Der pII.
XX
OS Synthetic.
XX
PN WO9308279-A.
XX
PD 29-APR-1993.
XX
PF 15-OCT-1992; 92WO-US08637.
XX
PR 16-OCT-1991; 91US-0777859.
PR 08-MAY-1992; 92US-0881396.
XX
PA (IMMU-) IMMULOGIC PHARM CORP.
XX
PI Garman RD, Greenstein JL, Kuo MC, Rogers BL;
XX
DR WPI; 1993-152472/18.
XX
PT Isolated peptide(s) of Dermatophagoides protein allergens - for
PT diagnosis and treatment of sensitivity to house dust mite
XX
PS Claim 44; Fig 3; 176pp; English.
XX
CC The peptide is one of a series of overlapping peptides synthesised by
CC standard techniques to cover the whole Dermatophagoides
CC pteronyssinus Der pII sequence. The T cell epitopes of the protein
CC were mapped by detection of the peptide's ability to stimulate T
CC cell activity. The peptides may be used for diagnosis and treatment
CC of sensitivity to house dust mite allergens. When administered to
CC house dust mite sensitive individuals, the peptides are capable of
CC modifying the allergic response to the allergens. The peptides may
CC be modified for e.g. increasing solubility, enhancing therapeutic or
CC preventive efficacy or stability.
CC See also AAR34686-700 and AAR36398-490.
XX
SQ Sequence 26 AA;

Query Match          50.9%; Score 84; DB 14; Length 26;
Best Local Similarity 63.6%; Pred. No. 3.7e-06;
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31
Db 3 VDVKDCANHEIKKVLVPGCHGS 24

RESULT 15
AAR51772
ID AAR51772 standard; Protein; 26 AA.
XX
AC AAR51772;
XX

```

```

DT 01-FEB-1995 (first entry)
XX
DE Der p II derived peptide, DP II-20(1-26).
XX
KW Group I; protein allergen; house dust mite; D. pteronyssinus; Der p I;
KW homology; D. farinae; Der f I; group II; Der p II; Der f II; T-cell;
KW epitopes; fusion peptides; antigenic fragments; substitution; deletion;
KW addition; chemical synthesis; chemical cleavage; recombinant techniques;
KW allergic response; immunoglobulin E; IgE; immunotherapy; anaphylaxis;
KW IgE-mediated responses; anergise; lymphokine secretion profile; modify;
KW T cell subpopulations; unresponsive; immune response; tolerance.
XX
OS Dermatophagoides pteronyssinus.
XX
PN ZA9302677-A.
XX
PD 26-JAN-1994.
XX
PF 16-APR-1993; 93ZA-0002677.
PR 16-APR-1993; 93ZA-0002677.
XX
PA (IMMU-) IMMULOGIC PHARM CORP.
XX
PI Garman RD, Greenstein JL, Kuo M, Rogers BL;
XX
DR WPI; 1994-126807/15.
XX
PT Isolated and/or modified peptides comprising T-cell epitopes - of
PT major protein allergens of genus Dermatophagoides, used to treat
PT or diagnose sensitivity to house dust mites
XX
PS Claim 28; Page 70; 154pp; English.
XX
CC The sequences given in AAR51731-841 represent T-cell epitopes derived
CC from the group I and II protein allergens from the house dust mite D.
CC farinae and D. pteronyssinus, Der f I, Der f II, Der p I and Der p II
CC respectively. The Der f II proteinsCC shows high homology having an
CC identity of 88%, with an identity of 81% between the two group I
CC proteins (see also AAR51727-30). Fusion peptides may be produced which
CC comprise at least two or these antigenic fragments. Each region of
CC these fusion peptides may be derived from the same, or different, mite
CC allergens. The antigenic fragments may be altered by substitution,
CC deletion or addition to enhance their antigenicity. These peptides may
CC be produced by chemical synthesis, chemical cleavage of the protein
CC allergen or by recombinant techniques. These peptides, or the fusion
CC peptides, when administered to a house dust mite sensitive individual,
CC are capable of modifying the allergic response of the individual to the
CC allergen. The peptides do not bind to immunoglobulin E (IgE), or bind
CC IgE to a lesser extent than the full length protein allergen. This
CC reduces the major complications of standard immunotherapy, which are
CC IgE-mediated responses such as anaphylaxis. Exposure of mite allergic
CC patients to these peptides may tolerate or anergise appropriate T cell
CC subpopulations such that they become unresponsive to mite allergens and
CC do not participate in mounting an immune response upon exposure.
CC Administration of the peptides may also modify the lymphokine secretion
CC profile as compared with exposure to the naturally occurring mite protein
CC allergen.
XX
SQ Sequence 26 AA;

Query Match          50.9%; Score 84; DB 15; Length 26;
Best Local Similarity 63.6%; Pred. No. 3.7e-06;
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31
Db 3 VDVKDCANHEIKKVLVPGCHGS 24

Search completed: December 4, 2002, 12:51:14
Job time : 16.6208 secs

```

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run On: December 4, 2002, 12:53:35 : Search time 2.84167 Seconds  
(without alignments)  
177.189 Million cell updates/sec

Title: US-09-362-731A-1  
Perfect score: 165  
Sequence: 1 QYKANSKFIGITELGGHEIKKVLVPGCHGS 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	ID	Description
1	84	50.9	129 10	US-09-949-889-4 Sequence 4, Appli
2	84	50.9	145 10	US-09-949-889-3 Sequence 3, Appli
3	81.5	49.4	28 10	US-09-848-834A-11 Sequence 11, Appl
4	78.5	47.6	29 10	US-09-732-754-1 Sequence 1, Appli
5	76	46.1	31 10	US-09-943-548-2 Sequence 2, Appli
6	76	46.1	31 10	US-09-983-019-5 Sequence 5, Appli
7	74.5	45.2	46 10	US-09-848-834A-19 Sequence 19, Appl
8	74	44.8	15 10	US-09-862-849-2 Sequence 2, Appli
9	74	44.8	16 10	US-09-848-834A-2 Sequence 2, Appli
10	74	44.8	31 10	US-09-848-834A-15 Sequence 15, Appl
11	74	44.8	50 10	US-09-943-548-8 Sequence 8, Appli
12	70	42.4	29 10	US-09-983-019-8 Sequence 8, Appli
13	70	42.4	29 10	US-09-983-019-9 Sequence 9, Appli
14	70	42.4	31 10	US-09-983-019-3 Sequence 3, Appli
15	70	42.4	31 10	US-09-983-019-6 Sequence 6, Appli
16	70	42.4	50 10	US-09-943-548-9 Sequence 9, Appli
17	63	38.2	22 10	US-09-860-793-7 Sequence 7, Appli
18	58.5	35.5	14 9	US-10-044-034-21 Sequence 21, Appl
19	54	32.7	28 10	US-09-864-761-47156 Sequence 47156, A

20 49 29.7 111 10 US-09-864-761-40073 Sequence 40073, A  
21 49 29.7 391 10 US-09-815-242-5655 Sequence 5655, Ap  
22 49 29.7 394 10 US-09-815-242-12267 Sequence 12267, A  
23 47 28.5 440 10 US-09-919-497-80 Sequence 80, Appl  
24 47 28.5 877 10 US-09-815-242-4901 Sequence 4901, Ap  
25 47 28.5 880 10 US-09-815-242-10491 Sequence 10491, A  
26 46.5 28.2 377 10 US-09-838-573-2 Sequence 2, Appli  
27 46 27.9 38 10 US-09-864-761-35665 Sequence 35665, A  
28 45 27.3 333 10 US-09-765-272-54 Sequence 54, Appl  
29 45 27.3 359 10 US-09-815-242-13169 Sequence 13169, A  
30 45 27.3 359 10 US-09-815-242-13593 Sequence 13593, A  
31 45 27.3 692 10 US-09-815-242-11568 Sequence 11568, A  
32 44.5 27.0 336 10 US-09-878-766A-20 Sequence 20, Appl  
33 44 26.7 64 10 US-09-925-300-1637 Sequence 1637, Ap  
34 44 26.7 176 10 US-09-764-869-1084 Sequence 1084, Ap  
35 44 26.7 615 10 US-09-940-101-2 Sequence 2, Appli  
36 44 26.7 1308 10 US-09-940-101-2 Sequence 2, Appli  
37 43.5 26.4 1080 10 US-09-119-855-2 Sequence 2, Appli  
38 43.5 26.4 1083 9 US-09-965-830-2 Sequence 2, Appli  
39 43 26.1 48 10 US-09-764-853-726 Sequence 726, App  
40 43 26.1 383 10 US-09-756-283A-27 Sequence 27, Appl  
41 43 26.1 1668 10 US-09-815-242-5654 Sequence 5654, Ap  
42 43 26.1 2397 10 US-09-815-242-12265 Sequence 12265, A  
43 42.5 25.8 197 10 US-09-935-390A-25 Sequence 25, Appl  
44 42.5 25.8 386 10 US-09-925-301-1277 Sequence 1277, Ap  
45 42 25.5 453 10 US-09-815-242-10263 Sequence 10263, A

## ALIGNMENTS

RESULT 1  
US-09-949-889-4  
; Sequence 4, Application US/09949889  
; Patent No. US20020054881A1  
; GENERAL INFORMATION:  
; APPLICANT: CONSIGLIO NAZIONALE DELLE RICERCHE  
; TITLE OF INVENTION: VARIANTS OF ALLERGENIC PROTEINS OF THE GROUP 2 OF  
; TITLE OF INVENTION: DERMATOPHAGOIDES  
; FILE REFERENCE: Cons Naz Ric  
; CURRENT APPLICATION NUMBER: US/09/949,889  
; CURRENT FILING DATE: 2001-09-12  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 129  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus  
US-09-949-889-4

Query Match 50.9%; Score 84; DB 10; Length 129;  
Best Local Similarity 63.6%; Pred. No. 7.2e-06;  
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31  
: : : |||||  
Db 3 VDKVDCANHEIKKVLVPGCHGS 24

RESULT 2  
US-09-949-889-3  
; Sequence 3, Application US/09949889  
; Patent No. US20020054881A1  
; GENERAL INFORMATION:  
; APPLICANT: CONSIGLIO NAZIONALE DELLE RICERCHE  
; TITLE OF INVENTION: VARIANTS OF ALLERGENIC PROTEINS OF THE GROUP 2 OF  
; TITLE OF INVENTION: DERMATOPHAGOIDES  
; FILE REFERENCE: Cons Naz Ric  
; CURRENT APPLICATION NUMBER: US/09/949,889  
; CURRENT FILING DATE: 2001-09-12  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3

```
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-09-949-889-3

Query Match      50.9%; Score 84; DB 10; Length 145;
Best Local Similarity 63.6%; Pred. No. 8.2e-06;
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 10 IGITELGGHEIKKVLVPGCHGS 31
Db 19 VDVKDCANHEIKKVLVPGCHGS 40

RESULT 3
US-09-848-834A-11
; Sequence 11, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphion Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 830-844 of the
; OTHER INFORMATION: Tetanus toxoid precursor (Tentoxylisin) linked by a spacer to am
; OTHER INFORMATION: ino acid sequence 2-10 of the GnRH hormone
; NAME/KEY: MOD.RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Amidated-glutamine
; NAME/KEY: MOD.RES
; LOCATION: (28)..(28)
; OTHER INFORMATION: Amidated-glycine or glycnamide
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(15)
; OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus Toxoid Precursor
; OTHER INFORMATION: (Tentoxylisin)
; NAME/KEY: PEPTIDE
; LOCATION: (16)..(19)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (20)..(28)
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-11

Query Match      49.4%; Score 81.5; DB 10; Length 28;
Best Local Similarity 61.3%; Pred. No. 3e-06;
Matches 19; Conservative 0; Mismatches 3; Indels 9; Gaps 1;

Qy 1 QYIKANSKFIGITELGGHEIKKVLVPGCHGS 31
Db 1 QYIKANSKFIGITELG-----PSLHWS 22

RESULT 4
US-09-732-754-1
; Sequence 1, Application US/09732754
; Patent No. US20020031523A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; APPLICANT: GRAS-MASSE, HELENE
; APPLICANT: BENMOHAMED, ELACHIR
; TITLE OF INVENTION: SYSTEMIC IMMUNE RESPONSE INDUCED BY MUCOSAL ADMINISTRATION OF LIH
; TITLE OF INVENTION: TAILED POLYPEPTIDES WITHOUT ADJUVANT
```

```
; FILE REFERENCE: 2008050555
; CURRENT APPLICATION NUMBER: US/09/732,754
; CURRENT FILING DATE: 2000-12-11
; PRIOR APPLICATION NUMBER: US 60/169,952
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Polypeptide
; NAME/KEY: LIPID
; LOCATION: (1)..(1)
; OTHER INFORMATION: N-epsilon Pam
US-09-732-754-1

Query Match      47.6%; Score 78.5; DB 10; Length 29;
Best Local Similarity 60.0%; Pred. No. 8.9e-06;
Matches 18; Conservative 2; Mismatches 5; Indels 5; Gaps 1;

Qy 1 QYIKANSKFIGITELGGHEIKKVLVPGCHG 30
Db 4 QYIKANSKFIGITERG-----RIKKEPVHG 28

RESULT 5
US-09-943-548-2
; Sequence 2, Application US/09943548
; Patent No. US20020042364A1
; GENERAL INFORMATION:
; APPLICANT: Rittershaus, Charles W.
; APPLICANT: Thomas, Lawrence J.
; TITLE OF INVENTION: MODULATION OF CHOLESTERYL ESTER TRANSFER PROTEIN (CETP) ACTIVI
; FILE REFERENCE: TCS-411.1P US-1; Tcs-411.1P US-2
; CURRENT APPLICATION NUMBER: US/09/943,548
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 08/432,483
; PRIOR FILING DATE: 1995-05-01
; PRIOR APPLICATION NUMBER: PCT/US96/06147
; PRIOR FILING DATE: 1996-05-01
; PRIOR APPLICATION NUMBER: 08/945,289
; PRIOR FILING DATE: 1997-10-17
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: vaccine peptide of the invention
US-09-943-548-2

Query Match      46.1%; Score 76; DB 10; Length 31;
Best Local Similarity 93.8%; Pred. No. 2.3e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITELG 16
Db 2 QYIKANSKFIGITEFG 17

RESULT 6
US-09-983-019-5
; Sequence 5, Application US/09983019
; Patent No. US20020146820A1
; GENERAL INFORMATION:
; APPLICANT: Diamond, Don J.
; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS P
; FILE REFERENCE: 1954-347US
; CURRENT APPLICATION NUMBER: US/09/983,019
; CURRENT FILING DATE: 2001-10-22
```

```

; NAME/KEY: misc_feature
; LOCATION: {}..{}
; OTHER INFORMATION: Human cytomegalovirus vaccine peptide
;
; NAME/KEY: LIPID
; LOCATION: (1)..(1)
; OTHER INFORMATION: palmitic acid
US-09-983-019-5

```

```

: APPLICANT: Gennady Goloobov
:
: TITLE OF INVENTION: Methods for Identifying Inducers and Inhibitors of
: Inducible Protein Expression in Cells
:
: FILE OF INVENTION: Antibodies, Compositions and Their Uses
:
: FILE REFERENCE: UNMC 63123 DIV
:
: CURRENT APPLICATION NUMBER: US/09/862,849
:
: CURRENT FILING DATE: 2001-08-29
:
: PRIOR APPLICATION NUMBER: US 09/046,373
:
: PRIOR FILING DATE: 1998-03-23
:
: NUMBER OF SEQ ID NOS: 10
:
: SOFTWARE: FastSeq for Windows version 3.0
:

```

```

; SOFTWARE: FASTSEQ FOR WINDOWS
; SEQ ID NO 2
; LENGTH: 15
; TYPE: PR1
; ORGANISM: Clostridium tetani
US-09-862-849-2

```

```

Query Match      44.48%; Score 74; DB 10; Length 15;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels

Qy 1 QYIKANSKEFIGITEL 15
    | | | | | | | | | | | | | | |
Db 1 QYIKANSKEFIGITEL 15

```

```

> RESULT 9
> US-09-848-834A-2
> Sequence 2, Application US/09848834A
> Patent No. US20020076416A1
> GENERAL INFORMATION:
> APPLICANT: Aptton Corporation
> TITLE OF INVENTION: Chimeric Peptide Immunogens
> FILE REFERENCE: 1102865-0047
> CURRENT APPLICATION NUMBER: US/09/848,834A
> CURRENT FILING DATE: 2001-05-04
> PRIOR APPLICATION NUMBER: 60/202,328
> PRIOR FILING DATE: 2000-05-05
> NUMBER OF SEQ ID NOS: 20
> SOFTWARE: PatentIn version 3.0
> SEQ ID NO 2
> LENGTH: 16
> TYPE: PPT
> ORGANISM: Tetanus bacillus
> FEATURE:
> NAME/KEY: PEPTIDE
> LOCATION: (1)..(16)
> OTHER INFORMATION: Amino acid sequence 829-844 of the Tetanus
> OTHER INFORMATION: Toxoid precursor (Tentoxylysin)
> US-09-848-834A-2

```

```

Query Match      44.8%; Score 74; DB 10; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels

Qy 1 QYIKANSKFIGITEL 15
    | | | | | | | | | |
Db 2 QYIKANSKFIGITEL 16
    | | | | | | | | | |

```

```
RESULT 10
US-09-848-834A-15
; Sequence 15, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aptton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848, 834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the GnRH hormone
; OTHER INFORMATION: RH hormone linked by a spacer to amino sequence 830-844 of the Tetanus toxin precursor (Tentoxylisin)
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(16)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (17)..(31)
; OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus toxin precursor
US-09-848-834A-15

Query Match 44.8%; Score 74; DB 10; Length 31;
Best Local Similarity 100.0%; Pred. No. 4.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15
Db 17 QYIKANSKFIGITEL 31

RESULT 11
US-09-943-548-8
; Sequence 8, Application US/09943548
; Patent No. US20020042364A1
; GENERAL INFORMATION:
; APPLICANT: Rittershaus, Charles W.
; APPLICANT: Thomas, Lawrence J.
; TITLE OF INVENTION: MODULATION OF CHOLESTERYL ESTER TRANSFER PROTEIN (CETP) ACTIVITY
; FILE REFERENCE: TCS-411.1P US-1; TCS-411.1P US-2
; CURRENT APPLICATION NUMBER: US/09/943,548
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 08/432,483
; PRIOR FILING DATE: 1995-05-01
; PRIOR APPLICATION NUMBER: PCT/US96/06147
; PRIOR FILING DATE: 1996-05-01
; PRIOR APPLICATION NUMBER: 08/945,289
; PRIOR FILING DATE: 1997-10-17
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: vaccine peptide of the invention
```

US-09-943-548-8

```
Query Match 44.8%; Score 74; DB 10; Length 50;
Best Local Similarity 100.0%; Pred. No. 8.1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 QYIKANSKFIGITEL 15
Db 2 QYIKANSKFIGITEL 16
```

RESULT 12

```
US-09-983-019-8
; Sequence 8, Application US/09983019
; Patent No. US20020146820A1
; GENERAL INFORMATION:
; APPLICANT: Diamond, Don J.
; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS P
; FILE REFERENCE: 1954-347US
; CURRENT APPLICATION NUMBER: US/09/983,019
; CURRENT FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/241,944
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: ( )..( )
; OTHER INFORMATION: Human cytomegalovirus vaccine peptide
US-09-983-019-8
```

```
Query Match 42.4%; Score 70; DB 10; Length 29;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 QYIKANSKFIGITE 14
Db 16 QYIKANSKFIGITE 29
```

RESULT 13

```
US-09-983-019-9
; Sequence 9, Application US/09983019
; Patent No. US20020146820A1
; GENERAL INFORMATION:
; APPLICANT: Diamond, Don J.
; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS P
; FILE REFERENCE: 1954-347US
; CURRENT APPLICATION NUMBER: US/09/983,019
; CURRENT FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/241,944
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: ( )..( )
; OTHER INFORMATION: Human cytomegalovirus vaccine peptide
US-09-983-019-9
```

```
Query Match 42.4%; Score 70; DB 10; Length 29;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 QYIKANSKFIGITE 14
```

```
Db 1 QYIKANSKFIGITE 14
|||||
RESULT 14
US-09-983-019-3
; Sequence 3, Application US/09983019
; Patent No. US20020146820A1
; GENERAL INFORMATION:
; APPLICANT: Diamond, Don J.
; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPTIOPES OF HUMAN CYTOMEGALOVIRUS PP15
; FILE REFERENCE: 1954-347US
; CURRENT APPLICATION NUMBER: US/09/983,019
; CURRENT FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/241,944
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (...)
; OTHER INFORMATION: Human cytomegalovirus vaccine peptide
; NAME/KEY: LIPID
; LOCATION: (1)..(1)
; OTHER INFORMATION: di-palmitic acid
US-09-983-019-3

Query Match 42.4%; Score 70; DB 10; Length 31;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITE 14
|||||
Db 4 QYIKANSKFIGITE 17

RESULT 15
US-09-983-019-6
; Sequence 6, Application US/09983019
; Patent No. US20020146820A1
; GENERAL INFORMATION:
; APPLICANT: Diamond, Don J.
; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPTIOPES OF HUMAN CYTOMEGALOVIRUS PP15
; FILE REFERENCE: 1954-347US
; CURRENT APPLICATION NUMBER: US/09/983,019
; CURRENT FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/241,944
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (...)
; OTHER INFORMATION: Human cytomegalovirus vaccine peptide
; NAME/KEY: LIPID
; LOCATION: (1)..(1)
; OTHER INFORMATION: palmitic acid
US-09-983-019-6

Query Match 42.4%; Score 70; DB 10; Length 31;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITE 14
|||||
```

Db 18 QYIKANSKFIGITE 31

Search completed: December 4, 2002, 13:05:28  
Job time : 3.84167 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 12:49:07 ; Search time 5.16667 seconds  
(without alignments)  
176.538 Million cell updates/sec

Title: US-09-362-731A-1  
Perfect score: 165  
Sequence: 1 QYIKANSKFIGITELGGHKKVLVPGCHGS 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A-COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B-COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A-COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B-COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92	55.8	47	1	US-08-446-692-35
2	92	55.8	47	2	US-08-488-351A-35
3	86	52.1	37	1	US-08-446-692-57
4	86	52.1	37	1	US-08-446-692-63
5	86	52.1	37	2	US-08-488-351A-57
6	86	52.1	37	2	US-08-488-351A-63
7	84	50.9	25	2	US-08-482-142-169
8	84	50.9	25	2	US-08-478-572-169
9	84	50.9	25	4	US-08-484-296-169
10	84	50.9	25	5	PCT-US95-04481-31
11	84	50.9	26	2	US-08-482-142-50
12	84	50.9	26	2	US-08-482-142-170
13	84	50.9	26	2	US-08-478-572-50
14	84	50.9	26	2	US-08-478-572-170
15	84	50.9	26	4	US-08-484-296-50
16	84	50.9	26	4	US-08-484-296-170
17	84	50.9	27	2	US-08-482-142-171
18	84	50.9	27	2	US-08-478-572-171
19	84	50.9	27	4	US-08-484-296-171
20	84	50.9	35	3	US-08-460-040-7
21	84	50.9	129	1	US-07-945-288-12
22	84	50.9	129	1	US-08-462-831-12
23	84	50.9	129	1	US-08-461-809-12
24	84	50.9	129	1	US-08-461-441-12
25	84	50.9	129	5	PCT-US93-08518-12
26	84	50.9	145	3	US-08-460-040-6
27	84	50.9	146	1	US-07-945-288-4

28	84	50.9	146	1	US-08-462-831-4	Sequence 4, Appli
29	84	50.9	146	1	US-08-461-809-4	Sequence 4, Appli
30	84	50.9	146	1	US-08-461-441-4	Sequence 4, Appli
31	84	50.9	146	2	US-08-482-142-4	Sequence 4, Appli
32	84	50.9	146	2	US-08-478-572-4	Sequence 4, Appli
33	84	50.9	146	4	US-08-484-296-4	Sequence 4, Appli
34	84	50.9	146	5	PCT-US93-08518-4	Sequence 4, Appli
35	83	50.3	26	2	US-08-482-142-52	Sequence 52, Appl
36	83	50.3	26	2	US-08-478-572-52	Sequence 52, Appl
37	83	50.3	26	4	US-08-484-296-52	Sequence 52, Appl
38	82	49.7	26	2	US-08-482-142-51	Sequence 51, Appl
39	82	49.7	26	2	US-08-478-572-51	Sequence 51, Appl
40	82	49.7	26	4	US-08-484-296-51	Sequence 51, Appl
41	79	47.9	16	2	US-08-482-142-59	Sequence 59, Appl
42	79	47.9	16	2	US-08-478-572-59	Sequence 59, Appl
43	79	47.9	16	4	US-08-484-296-59	Sequence 59, Appl
44	79	47.9	25	2	US-08-482-142-42	Sequence 42, Appl
45	79	47.9	25	2	US-08-478-572-42	Sequence 42, Appl

ALIGNMENTS

RESULT 1  
US-08-446-692-35  
; Sequence 35, Application US/08446692  
; Patent No. 5759551  
; GENERAL INFORMATION:  
; APPLICANT: Ladd, Anna  
; APPLICANT: Wang, Chang Yi  
; APPLICANT: Zamb, Timothy  
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs  
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines  
; NUMBER OF SEQUENCES: 114  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Maria C.H. Lin  
; STREET: 345 Park Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: US  
; ZIP: 10154-0053  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/446,692  
; FILING DATE: 7-JUN-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maria C.H. Lin  
; REGISTRATION NUMBER: 29,323  
; REFERENCE/DOCKET NUMBER: 1151-4146 US2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)415-8745  
; TELEFAX: (516)751-6849  
; INFORMATION FOR SEQ ID NO: 35:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 47 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-446-692-35

Query Match 55.8%; Score 92; DB 1; Length 47;  
Best Local Similarity 74.1%; Pred. No. 5e-07;  
Matches 20; Conservative 0; Mismatches 7; Indels 0; Caps 0;  
QY 1 QYIKANSKFIGITELGGHKKVLVPG 27  
|||||  
DB 21 QYIKANSKFIGITELGGHWSYGLRPG 47



```
RESULT 2
US-08-488-351A-35
; Sequence 35, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488.351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-35

Query Match 55.8%; Score 92; DB 2; Length 47;
Best Local Similarity 74.1%; Pred. No. 5e-07; Indels 7; Gaps 0;
Matches 20; Conservative 0; Mismatches 7;

Oy 1 QYIKANSKFIGITELGGHKKVLVPG 27
Db 21 QYIKANSKFIGITELGGHWSYGLRPG 47

RESULT 3
US-08-446-692-57
; Sequence 57, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-446-692-57

Query Match 52.1%; Score 86; DB 1; Length 37;
Best Local Similarity 100.0%; Pred. No. 3.1e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 QYIKANSKFIGITELGG 17
Db 21 QYIKANSKFIGITELGG 37

RESULT 4
US-08-446-692-63
; Sequence 63, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-446-692-57

Query Match 52.1%; Score 86; DB 1; Length 37;
Best Local Similarity 100.0%; Pred. No. 3.1e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 QYIKANSKFIGITELGG 17
Db 21 QYIKANSKFIGITELGG 37
```

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)415-8745  
TELEFAX: (516)751-6849  
INFORMATION FOR SEQ ID NO: 63:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 37 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-446-692-63

Query Match 52.1%; Score 86; DB 1; Length 37;  
Best Local Similarity 100.0%; Pred. No. 3.le-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITELGG 17  
|||||

Db 5 QYIKANSKFIGITELGG 21

## RESULT 5

US-08-488-351A-57  
Sequence 57, Application US/08488351A  
Patent No. 5843446

## GENERAL INFORMATION:

APPLICANT: Ladd, Anna  
APPLICANT: Wang, Chang Yi  
APPLICANT: Zamb, Timothy

TITLE OF INVENTION: Immunogenic LHRH peptide constructs

TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines

NUMBER OF SEQUENCES: 114

CORRESPONDENCE ADDRESS:

ADDRESSEE: Maria C.H. Lin

STREET: 345 Park Avenue

CITY: New York

STATE: NY

COUNTRY: US

ZIP: 10154-0053

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/488,351A

FILING DATE: 7-JUN-1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/446,692

FILING DATE: 7-JUN-1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/229,275

FILING DATE: 14-APR-1994

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/057,166

FILING DATE: 27-APR-1992

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Maria C.H. Lin

REGISTRATION NUMBER: 29,323

REFERENCE/DOCKET NUMBER: 1151-4146 US2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212)415-8745

TELEFAX: (516)751-6849

INFORMATION FOR SEQ ID NO: 57:

SEQUENCE CHARACTERISTICS:

LENGTH: 37 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-488-351A-57

Query Match 52.1%; Score 86; DB 2; Length 37;  
Best Local Similarity 100.0%; Pred. No. 3.le-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITELGG 17

|||||

Db 21 QYIKANSKFIGITELGG 37

## RESULT 6

US-08-488-351A-63  
Sequence 63, Application US/08488351A  
Patent No. 5843446

## GENERAL INFORMATION:

APPLICANT: Ladd, Anna

APPLICANT: Wang, Chang Yi

APPLICANT: Zamb, Timothy

TITLE OF INVENTION: Immunogenic LHRH peptide constructs

TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines

NUMBER OF SEQUENCES: 114

CORRESPONDENCE ADDRESS:

ADDRESSEE: Maria C.H. Lin

STREET: 345 Park Avenue

CITY: New York

STATE: NY

COUNTRY: US

ZIP: 10154-0053

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/488,351A

FILING DATE: 7-JUN-1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/446,692

FILING DATE: 7-JUN-1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/229,275

FILING DATE: 14-APR-1994

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/057,166

FILING DATE: 27-APR-1992

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Maria C.H. Lin

REGISTRATION NUMBER: 29,323

REFERENCE/DOCKET NUMBER: 1151-4146 US2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212)415-8745

TELEFAX: (516)751-6849

INFORMATION FOR SEQ ID NO: 63:

SEQUENCE CHARACTERISTICS:

LENGTH: 37 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-488-351A-63

Query Match 52.1%; Score 86; DB 2; Length 37;  
Best Local Similarity 100.0%; Pred. No. 3.le-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITELGG 17

|||||

Db 5 QYIKANSKFIGITELGG 21

## RESULT 7

US-08-482-142-169  
; Sequence 169, Application US/08482142  
; Patent No. 5820862  
; GENERAL INFORMATION:  
; APPLICANT: Garman, Richard  
; APPLICANT: Greenstein, Julia  
; APPLICANT: Kuo, Mei-chang  
; APPLICANT: Rogers, Bruce  
; APPLICANT: Franzen, Henry  
; APPLICANT: Chen, Xian  
; APPLICANT: Evans, Sean  
; APPLICANT: Shaked, Ze'ev  
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
; FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
; NUMBER OF SEQUENCES: 207  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
; STREET: 610 LINCOLN STREET  
; CITY: WALTHAM  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02154

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/482,142  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/445,307  
; FILING DATE: 07 June 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CRAIG, ANNE I.  
; REGISTRATION NUMBER: 32,976  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 466-6000  
; TELEFAX: (617) 466-6040  
; INFORMATION FOR SEQ ID NO: 169:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 25 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: internal  
US-08-482-142-169

Query Match 50.9%; Score 84; DB 2; Length 25;  
Best Local Similarity 63.6%; Pred. No. 4.1e-06;  
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31  
; : : : |||||  
Db 3 VDVKDCANHEIKKVLVPGCHGS 24

## RESULT 8

US-08-478-572-169  
; Sequence 169, Application US/08478572  
; Patent No. 5968526  
; GENERAL INFORMATION:  
; APPLICANT: Garman, Richard  
; APPLICANT: Greenstein, Julia  
; APPLICANT: Kuo, Mei-chang  
; APPLICANT: Rogers, Bruce  
; APPLICANT: Franzen, Henry  
; APPLICANT: Chen, Xian  
; APPLICANT: Evans, Sean  
; APPLICANT: Shaked, Ze'ev

; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
; FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
; NUMBER OF SEQUENCES: 207  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
; STREET: 610 LINCOLN STREET  
; CITY: WALTHAM  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/478,572  
; FILING DATE: 07-June-1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/445,307  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CRAIG, ANNE I.  
; REGISTRATION NUMBER: 32,976  
; REFERENCE/DOCKET NUMBER: 017.605  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 466-6000  
; TELEFAX: (617) 466-6040  
; INFORMATION FOR SEQ ID NO: 169:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 25 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: internal  
US-08-478-572-169

Query Match 50.9%; Score 84; DB 2; Length 25;  
Best Local Similarity 63.6%; Pred. No. 4.1e-06;  
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31  
; : : : |||||  
Db 3 VDVKDCANHEIKKVLVPGCHGS 24

## RESULT 9

US-08-484-296-169  
; Sequence 169, Application US/08484296  
; Patent No. 6268491  
; GENERAL INFORMATION:  
; APPLICANT: Garman, Richard  
; APPLICANT: Greenstein, Julia  
; APPLICANT: Kuo, Mei-chang  
; APPLICANT: Rogers, Bruce  
; APPLICANT: Franzen, Henry  
; APPLICANT: Chen, Xian  
; APPLICANT: Evans, Sean  
; APPLICANT: Shaked, Ze'ev

; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
; FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
; NUMBER OF SEQUENCES: 207  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
; STREET: 610 LINCOLN STREET  
; CITY: WALTHAM  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,296  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/445,307  
FILING DATE: 07 June 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: CRAIG, ANNE I.  
REGISTRATION NUMBER: 32,976  
REFERENCE/DOCKET NUMBER: 017.6US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 466-6000  
TELEFAX: (617) 466-6040  
INFORMATION FOR SEQ ID NO: 169:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 25 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-484-296-169

Query Match 50.9%; Score 84; DB 4; Length 25;  
Best Local Similarity 63.6%; Pred. No. 4.1e-06;  
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31  
Db 3 VDVKDCANHEIKKVLVPGCHGS 24

## RESULT 10

PCT-US95-04481-31  
Sequence 31, Application PC/TUS9504481  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: Pharmaceutical Peptide Formulations For Treatment of Dust Mit  
NUMBER OF SEQUENCES: 54  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04481  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/227,772  
FILING DATE: April 14, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Vanstone, Darlene A.  
REGISTRATION NUMBER: 35,279  
REFERENCE/DOCKET NUMBER: 017.5 PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 466-6000  
TELEFAX: (617) 466-6040  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 25 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US95-04481-31

Query Match 50.9%; Score 84; DB 5; Length 25;  
Best Local Similarity 63.6%; Pred. No. 4.1e-06;  
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31  
Db 3 VDVKDCANHEIKKVLVPGCHGS 24  
RESULT 11  
US-08-482-142-50  
Sequence 50, Application US/08482142  
Patent No. 5820862  
GENERAL INFORMATION:  
APPLICANT: Garman, Richard  
APPLICANT: Greenstein, Julia  
APPLICANT: Kuo, Mei-chang  
APPLICANT: Rogers, Bruce  
APPLICANT: Franzen, Henry  
APPLICANT: Chen, Xian  
APPLICANT: Evans, Sean  
APPLICANT: Shaked, Ze'ev  
TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
NUMBER OF SEQUENCES: 207  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
STREET: 610 LINCOLN STREET  
CITY: WALTHAM  
STATE: MA  
COUNTRY: USA  
ZIP: 02154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,142  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/445,307  
FILING DATE: 07 June 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: CRAIG, ANNE I.  
REGISTRATION NUMBER: 32,976  
REFERENCE/DOCKET NUMBER: 017.6US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 466-6000  
TELEFAX: (617) 466-6040  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 26 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal  
US-08-482-142-50

Query Match 50.9%; Score 84; DB 2; Length 26;  
Best Local Similarity 63.6%; Pred. No. 4.2e-06;  
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31  
Db 3 VDVKDCANHEIKKVLVPGCHGS 24

## RESULT 12

US-08-482-142-170  
Sequence 170, Application US/08482142  
Patent No. 5820862  
GENERAL INFORMATION:  
APPLICANT: Garman, Richard  
APPLICANT: Greenstein, Julia  
APPLICANT: Kuo, Mei-chang

```
; APPLICANT: Rogers, Bruce
; APPLICANT: Franzen, Henry
; APPLICANT: Chen, Xian
; APPLICANT: Evans, Sean
; APPLICANT: Shaked, Ze'ev
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS
; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,142
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,307
; FILING DATE: 07 June 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAIG, ANNE I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: 017.605
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6040
; INFORMATION FOR SEQ ID NO: 170:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-08-482-142-170

Query Match 50.9%; Score 84; DB 2; Length 26;
Best Local Similarity 63.6%; Pred. No. 4.2e-06;
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31
: : : |||||
Db 3 VDVKDCANHEIKKVLVPGCHGS 24

RESULT 13
US-08-478-572-50
; Sequence 50, Application us/08478572
; Patent No. 5968526
; GENERAL INFORMATION:
; APPLICANT: Garman, Richard
; APPLICANT: Greenstein, Julia
; APPLICANT: Kuo, Mei-chang
; APPLICANT: Rogers, Bruce
; APPLICANT: Franzen, Henry
; APPLICANT: Chen, Xian
; APPLICANT: Evans, Sean
; APPLICANT: Shaked, Ze'ev
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS
; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,572
; FILING DATE: 07-June-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/445,307
```

```
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,572
; FILING DATE: 07-June-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/445,307
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAIG, ANNE I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: 017.605
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6040
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; US-08-478-572-50

Query Match 50.9%; Score 84; DB 2; Length 26;
Best Local Similarity 63.6%; Pred. No. 4.2e-06;
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 10 IGITELGGHEIKKVLVPGCHGS 31
: : : |||||
Db 3 VDVKDCANHEIKKVLVPGCHGS 24

RESULT 14
US-08-478-572-170
; Sequence 170, Application us/08478572
; Patent No. 5968526
; GENERAL INFORMATION:
; APPLICANT: Garman, Richard
; APPLICANT: Greenstein, Julia
; APPLICANT: Kuo, Mei-chang
; APPLICANT: Rogers, Bruce
; APPLICANT: Franzen, Henry
; APPLICANT: Chen, Xian
; APPLICANT: Evans, Sean
; APPLICANT: Shaked, Ze'ev
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS
; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,572
; FILING DATE: 07-June-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/445,307
```

FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: CRAIG, ANNE I.  
REGISTRATION NUMBER: 32,976  
REFERENCE/DOCKET NUMBER: 017.605  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 466-6000  
TELEFAX: (617) 466-6040  
INFORMATION FOR SEQ ID NO: 170:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 26 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-478-572-170

Query Match 50.9%; Score 84; DB 2; Length 26;  
Best Local Similarity 63.6%; Pred. No. 4.2e-06;  
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
Qy 10 IGITELGGHEIKKVLVPGCHGS 31  
Db 3 VDVKDCANHEIKKVLVPGCHGS 24

RESULT 15  
US-08-484-296-50  
Sequence 50, Application US/08484296  
Patent No. 6268491  
GENERAL INFORMATION:  
APPLICANT: Garman, Richard  
APPLICANT: Greenstein, Julia  
APPLICANT: Kuo, Mei-chang  
APPLICANT: Rogers, Bruce  
APPLICANT: Franzen, Henry  
APPLICANT: Chen, Xian  
APPLICANT: Evans, Sean  
APPLICANT: Shaked, Ze'ev  
TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
NUMBER OF SEQUENCES: 207  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
STREET: 610 LINCOLN STREET  
CITY: WALTHAM  
STATE: MA  
COUNTRY: USA  
ZIP: 02154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,296  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/445,307  
FILING DATE: 07 June 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: CRAIG, ANNE I.  
REGISTRATION NUMBER: 32,976  
REFERENCE/DOCKET NUMBER: 017.605  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 466-6000  
TELEFAX: (617) 466-6040  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 26 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal  
US-08-484-296-50  
Query Match 50.9%; Score 84; DB 4; Length 26;  
Best Local Similarity 63.6%; Pred. No. 4.2e-06;  
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
Qy 10 IGITELGGHEIKKVLVPGCHGS 31  
Db 3 VDVKDCANHEIKKVLVPGCHGS 24  
Search completed: December 4, 2002, 12:55:07  
Job time : 6.16667 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: December 4, 2002, 12:49:07 : Search time 26.8292 Seconds  
(without alignments)  
490.899 Million cell updates/sec

Title: US-09-362-731A-3

Perfect score: 793

Sequence: 1 DQYIKANSKFTGIGELGQY.....FGGCHGSEPCIIHRGKPFSSR 137

Scoring table: BIOSUM62

Gapop 10:0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	101	12.7	1571	2 T14155	zinc finger protei
2	99	12.5	146	2 A60381	allergen Der p II
3	97	12.2	129	2 A61501	allergen Der f II
4	97	12.2	129	2 J00394	allergen Der f II
5	97	12.2	138	2 B61241	allergen Der f II
6	97	12.2	138	2 A61241	allergen Der f II
7	90	11.3	1700	2 S08167	Balbani ring 3 pr
8	88	11.1	615	2 S06546	finger protein (cl
9	85	10.7	169	2 S18946	ultra high-sulfur
10	85	10.7	232	2 T1524	hypothetical prote
11	84	10.6	230	2 A38346	ultra-high-sulfur
12	82.5	10.4	196	2 A81004	peptidylprolyl iso
13	82.5	10.4	1378	2 T30173	zinc finger protei
14	82	10.3	102	2 A72507	hypothetical prote
15	81.5	10.3	1584	2 T22674	hypothetical prote
16	81	10.2	846	2 H70599	hypothetical prote
17	80.5	10.2	197	2 S06564	finger protein (cl
18	80.5	10.2	1797	2 A55677	laminin beta-2 cha
19	80	10.1	361	2 T14460	oleosin homolog qb
20	80	10.1	1523	2 T13953	MEGF5 protein - ra
21	79	10.0	63	2 S25772	testis-specific pr
22	79	10.0	351	1 O2ZQKU	circumsporozoite p
23	79	10.0	769	2 A41029	integrin beta-8 ch
24	78	9.8	1315	1 BTCLTN	tentoxilysin (EC 3
25	78	9.8	1820	2 A55494	latent transformin
26	78	9.8	2101	2 S27245	insulin receptor (
27	78	9.8	2148	1 A56081	insulin receptor (
28	77.5	9.8	471	2 T33997	hypothetical prote
29	77.5	9.8	1801	1 MMRTS	laminin beta-2 cha

30 76.5 9.6 3871 2 T22812 hypothetical prote  
31 76 9.6 261 2 S70006 finger protein zfo  
32 76 9.6 262 2 S56100 outer dense fiber  
33 76 9.6 313 2 T04776 hypothetical prote  
34 76 9.6 419 2 T14448 oleosin homolog pc  
35 76 9.6 435 2 S00833 finger protein (cl  
36 76 9.6 1487 2 G96827 protein F20817.10  
37 75.5 9.5 373 2 B96798 extensin I imported  
38 75.5 9.5 646 2 H96665 protein F22C12.10  
39 75.5 9.5 788 2 B41029 integrin beta-8 ch  
40 75 9.5 284 2 S71227 extensin 1 - Arabi  
41 75 9.5 380 2 T14447 oleosin homolog pc  
42 75 9.5 955 2 A45441 thrombospondin 4 -  
43 74.5 9.4 1119 2 A88481 protein Cl6A3.6 [i  
44 74.5 9.4 1798 2 S53869 laminin beta-2 cha  
45 74.5 9.4 1964 2 T09059 notch4 - mouse

## ALIGNMENTS

## RESULT 1

T14155

zinc finger protein Peg3 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999

C:Accession: T14155

R:Kuroiwa, Y.; Kaneko-Ishino, T.; Kagitani, F.; Kohda, T.; Li, L.L.; Tada, M.; Suzuki  
Nature Genet. 12, 186-190, 1996

A:Title: Peg3 imprinted gene on proximal chromosome 7 encodes for a zinc finger prote

A:Reference number: Z17892; MUID:96154192; PMID:8563758

A:Accession: T14155

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1571 &lt;KUR&gt;

A:Cross-references: EMBL:AF038939; NID:g2791677; PID:g2791678; PIDN:AAB96922.1

C:Genetics:

A:Gene: Peg3

A:Map position: 7

C:Keywords: zinc finger

Query Match 12.7%; Score 101; DB 2; Length 1571;  
Best Local Similarity 33.3%; Pred. No. 0.049;  
Matches 40; Conservative 8; Mismatches 26; Indels 46; Gaps 12;

QY 37 HGSEPCIIHRGKPF-----GCHGSEP-----CII-----HRGKPFSSCHGSEPCII 78

Db 948 HGQK---IHDKETPGKPSGKEPHGDEPQDKPQKPLQEMRSEEPHDDKP---HGQEP--- 997

QY 79 HRGKPFSGCHGSPFCIIHRGKPFSSCHGSEPCIIHRGKPFPG-GCHGSEPCIIHRGKPFSSR 137

Db 998 HDDKP---HGQEP---HDDKP---HGQEP---HGDEPHGQEPHGDEP---HDKEPIDQ 1040

## RESULT 2

A60381

allergen Der p II precursor - house-dust mite (Dermatophagoides pteronyssinus)

C:Species: Dermatophagoides pteronyssinus

C:Date: 03-Mar-1993 #sequence\_revision 03-Mar-1993 #text\_change 13-Sep-1998

C:Accession: A60381

R:Chua, K.Y.; Doyle, C.R.; Simpson, R.J.; Turner, K.J.; Stewart, G.A.; Thomas, W.R.  
Int. Arch. Allergy Appl. Immunol. 91, 118-123, 1990

A:Title: Isolation of cDNA coding for the major mite allergen Der p II by IgE plaque

A:Reference number: A60381; MUID:90256301; PMID:2341191

A:Accession: A60381

A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-146 &lt;CHU&gt;

C:Superfamily: allergen Der p II

F:1-17/Domain: signal sequence #status predicted &lt;SIG&gt;

F:18-146/Product: allergen Der p II #status predicted &lt;MAT&gt;

Query Match 12.5%; Score 99; DB 2; Length 146;

Best Local Similarity 100.0%; Pred. No. 0.0077;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 GCHGSEPCIIHRGKPF 67  
|||||  
DB 37 GCHGSEPCIIHRGKPF 52

RESULT 3  
A61501  
allergen Der f II - house-dust mite (Dermatophagoides farinae) (fragment)  
C:Species: Dermatophagoides farinae  
C:Date: 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 13-Sep-1998  
C:Accession: A61501  
R:Trudinger, M.; Chua, K.Y.; Thomas, W.R.  
Clin. Exp. Allergy 21, 33-37, 1991  
A:Title: cDNA encoding the major mite allergen Der f II.  
A:Reference number: A61501; MUID:91215495; PMID:2021876  
A:Accession: A61501  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-129 <TRU>  
C:Superfamily: allergen Der p II

Query Match 12.2%; Score 97; DB 2; Length 129;  
Best Local Similarity 88.2%; Pred. No. 0.011;  
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 52 GCHGSEPCIIHRGKPF 68  
|||||  
DB 20 GCHGSDPCIIHRGKPF 36

RESULT 4  
JU0394  
allergen Der f II (pFL2) - house-dust mite (Dermatophagoides farinae)  
C:Species: Dermatophagoides farinae  
C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 17-Mar-1999  
C:Accession: JU0394  
R:Yuuki, T.; Okumura, Y.; Ando, T.; Yamakawa, H.; Suko, M.; Haida, M.; Okudaira  
Agric. Biol. Chem. 55, 1233-1238, 1991  
A:Title: Cloning and expression of cDNA coding for the major house dust mite al  
A:Reference number: PS0417; MUID:91291341; PMID:1368682  
A:Accession: JU0394  
A:Molecule type: mRNA  
A:Residues: 1-129 <YUU>  
C:Superfamily: allergen Der p II

Query Match 12.2%; Score 97; DB 2; Length 129;  
Best Local Similarity 88.2%; Pred. No. 0.011;  
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 52 GCHGSEPCIIHRGKPF 68  
|||||  
DB 20 GCHGSDPCIIHRGKPF 36

RESULT 5  
B61241  
allergen Der f II precursor - house-dust mite (Dermatophagoides farinae) (fragm  
C:Species: Dermatophagoides farinae  
C:Date: 12-May-1994 #sequence\_revision 27-Jun-1994 #text\_change 13-Sep-1998  
C:Accession: B61241; JU0395  
R:Yuuki, T.; Okumura, Y.; Ando, T.; Yamakawa, H.; Suko, M.; Haida, M.; Dohi, M.  
Int. Arch. Allergy Appl. Immunol. 94, 354-356, 1991  
A:Title: Synthesis of biologically active recombinant Der f II.  
A:Reference number: A61241; MUID:52040281; PMID:1937898  
A:Accession: B61241  
A:Molecule type: mRNA  
A:Residues: 1-138 <YUU>  
C:Superfamily: allergen Der p II  
F:1-9/Domain: signal sequence (fragment) #status predicted <SIG>  
F:10-138/Product: allergen Der f II #status predicted <MAT>



```

Query Match          10.7%; Score 85; DB 2; Length 232;
Best Local Similarity 35.6%; Pred. No. 0.28;
Matches 36; Conservative 6; Mismatches 35; Indels 24; Gaps 10;

QY 45 HRGKPF-GGCHGSEPCIHRGKPF-SSCHGSEPCIHRGKPF-GGCHGSEPCI--IHRGK 99
      |||: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 81 HRGETLGRGETHRGE---THRGTEFRGETHRGE---THRGTEFRGETHRGETFCGETHRGE 134

QY 100 PF-SSCHGSEPCIHRGKPF-----GGCHGSEPCIHRGK 133
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 135 TFRSETHRGE---THRGTEFRGETHRGEAHRGE---AHRGE 169

RESULT 11
A38346
ultra-high-sulfur keratin 1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 28-Jun-1991 #sequence_revision 28-Jun-1991 #text_change 24-Sep-1999
C:Accession: A38346
J. Biol. Chem. 265, 21375-21380, 1990
A:Title: Serine-rich ultra high sulfur protein gene expression in murine hair a
A:Reference number: A38346; M01D:91065960; PMID:2250030
A:Accession: A38346
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-230 <WOO>
A:Cross-references: GB:M37759; NID:g200961; PIDN:AAA40106.1; PID:g200962
C:Superfamily: ultra-high-sulfur keratin

```



Db 1068 ESSDCHINGHCVINEHGAGEYICQCLPGFSGDGFINCRGADQC--NPSNP-SACYQNAHC 1124  
Qy 77 I-----IHRGKPFGGCHG-SEPCIIHRGKPFSSCHGSEPCIIHRGKPFGGCHGSEPCII 129  
Db 1125 VYDAILNAHACKVDGFKDGTSCVPY--APATNCN-LEPRI-----CHANAQOCVM 1172  
Qy 130 H 130  
Db 1173 H 1173

Search completed: December 4, 2002, 12:54:22  
Job time : 27.8292 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	100	12.6	215	1	STRD_MOUSE	Q60924	mus musculus
2	99	12.5	146	1	ALL2_DERPT	P49278	dermatophag
3	97	12.2	146	1	ALL2_DERFA	Q00855	chironomus
4	90	11.3	1700	1	BAR3_CHITE	Q00376	chironomus
5	88	11.1	898	1	Z071_XENLA	P18751	xenopus lae
6	88	10.8	488	1	ZF92_MOUSE	Q62396	mus musculus
7	86	10.8	677	1	SP87_DICDI	P54643	dictyosteli
8	85.5	10.8	153	1	CORB_MOUSE	Q62267	mus musculus
9	85	10.7	169	1	KRUA_HUMAN	P26371	homo sapien
10	80.5	10.2	197	1	ZG17_XENLA	P18713	xenopus lae
11	79	10.0	63	1	M84A_DROME	Q01642	drosophila
12	79	10.0	351	1	CSP_PLAKU	P04922	plasmodium
13	79	10.0	769	1	ITB8_HUMAN	P26012	homo sapien
14	78	9.8	1314	1	TEFX_CLOFE	P04958	clostridium
15	78	9.8	2146	1	INSR_DROME	P05208	drosophila
16	77.5	9.8	1801	1	LMB2_RAT	P15800	rattus norv
17	77.5	9.8	4590	1	FATH_HUMAN	Q14517	homo sapien
18	76	9.6	262	1	ODFP_PIG	Q29077	sus scrofa
19	76	9.6	370	1	DNAJ_ERYRH	Q05646	erysipeloth
20	76	9.6	435	1	Z022_XENLA	P18745	xenopus lae
21	75.5	9.5	373	1	EXT1_ARATH	Q38913	arabidopsis
22	75.5	9.5	768	1	ITB8_RABIT	P26013	oryctolagus
23	75	9.5	145	1	ALL2_EURMA	Q9tzz2	euroglyphus
24	75	9.5	446	1	ZN70_HUMAN	Q9uc06	homo sapien
25	75	9.5	955	1	TSP4_XENLA	Q06441	xenopus lae
26	74.5	9.4	535	1	Z257_HUMAN	Q9y2q1	homo sapien
27	74.5	9.4	1654	1	PCPB_HUMAN	Q094913	homo sapien
28	74.5	9.4	1798	1	LMB2_HUMAN	P55268	homo sapien
29	74.5	9.4	1964	1	NTC4_MOUSE	P31695	mus musculus
30	74	9.3	262	1	ODFP_BOVIN	Q29438	bos taurus
31	73.5	9.3	474	1	Z141_HUMAN	Q15928	homo sapien
32	73.5	9.3	1387	1	TROP_HUMAN	Q12816	homo sapien
33	73	9.2	56	1	M87F_DROME	P08175	drosophila

```

RESULT 2
ALL2_DERPT
ID ALL2_DERPT STANDARD; PRT; 146 AA.
AC P49278;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mite group 2 allergen Der p 2 precursor (Der p II) (DPX).
GN DERP2.
OS Dermatophagoides pteronyssinus (House-dust mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcotiformes; Astigmata; Analgoidea; Pyroglyphidae;
OC Dermatophagoides.
OX NCBI_TaxID=6956;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90256301; PubMed=2341191;
RA Chua K.Y., Doyle C.R., Simpson R.J., Turner K.J., Stewart G.A.,
RA Thomas W.R.;
RT "Isolation of cDNA coding for the major mite allergen Der p II by IgE
RT plaque immunoassay.";
RL Int. Arch. Allergy Appl. Immunol. 91:118-123(1990).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANTS.
RX MEDLINE=21290932; PubMed=11398075;
RA Smith W.A., Hales B.J., Jarnicki A.G., Thomas W.R.;
RT "Allergens of wild house dust mites: environmental Der p 1 and Der p 2
RT sequence polymorphisms.";
RL J. Allergy Clin. Immunol. 107:985-992(2001).
RN [3]
RP PARTIAL SEQUENCE OF 18-57.
RX MEDLINE=89278484; PubMed=2732406;
RA Heymann P.W., Chapman M.D., Aalberse R.C., Fox J.W.,
RA Platts-Mills T.A.;
RT "Antigenic and structural analysis of group II allergens (Der f II
RT and Der p II) from house dust mites (Dermatophagoides spp).";
RL J. Allergy Clin. Immunol. 83:1055-1067(1989).
RN [4]
RP STRUCTURE BY NMR.
RX MEDLINE=98409423; PubMed=9737847;
RA Mueller G.A., Benjamin D.C., Rule G.S.;
RT "Tertiary structure of the major house dust mite allergen Der p 2:
RT sequential and structural homologies.";
RL Biochemistry 37:12707-12714(1998)
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE NPC2 FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; AF276239; AAF86462.1; -.
DR PDB; 1A9V; 14-OCT-98.
DR InterPro; IPR003172; El_Derp2_Derp2.
DR Pfam; PF02221; El_Derp2_Derp2; 1.
KW Allergen; Signal; 3D-structure; Polymorphism.
FT SIGNAL 1 17
FT CHAIN 18 146 MITE GROUP 2 ALLERGEN DER P 2.
FT DISULFID 25 136
FT DISULFID 38 44
FT DISULFID 90 95
FT VARIANT 39 39
FT VARIANT 40 40
FT VARIANT 44 44
FT VARIANT 47 47
FT VARIANT 49 49
FT VARIANT 56 56
H -> A.
G -> L.
C -> N.
H -> S.
G -> T.
A -> Y.

```

```

FT VARIANT 57 57
FT VARIANT 61 61
FT VARIANT 64 64
FT VARIANT 75 75
FT VARIANT 78 78
FT VARIANT 81 81
FT VARIANT 95 95
FT VARIANT 98 98
FT VARIANT 108 108
FT VARIANT 111 111
FT VARIANT 114 114
FT VARIANT 115 115
FT VARIANT 116 116
FT VARIANT 118 118
FT VARIANT 127 127
FT VARIANT 128 128
FT VARIANT 131 131
FT VARIANT 133 133
FT VARIANT 144 144
SQ SEQUENCE 146 AA; 15999 MW; 591B2FA7FD26D3AF CRC64;

Query Match 12.5%; Score 99; DB 1; Length 146;
Best Local Similarity 100.0%; Pred. NO. 0.0032;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 GCHGSEPCIHRGKPF 67
DB 37 GCHGSEPCIHRGKPF 52
IIIIIIIIIIIIIIIIII

RESULT 3
ALL2_DERFA
ID ALL2_DERFA STANDARD; PRT; 146 AA.
AC Q00855; P39672; Q26359;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mite group 2 allergen Der f 2 precursor (Der f II).
GN DERF2.
OS Dermatophagoides farinae (House-dust mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcotiformes; Astigmata; Analgoidea; Pyroglyphidae;
OC Dermatophagoides.
OX NCBI_TaxID=6954;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=91291341; PubMed=1368682;
RA Yuuki T., Okumura Y., Ando T., Yamakawa H., Suko M., Haida M.,
RA Okudaira H.;
RT "Cloning and expression of cDNA coding for the major house dust mite
RT allergen Der f II in Escherichia coli.";
RL Agric. Biol. Chem. 55:1233-1238(1991).
RN [2]
RP SEQUENCE OF 4-146 FROM N.A.
RX MEDLINE=94256850; PubMed=8198452;
RA Okuhira H.;
RT "Molecular biology of mite antigens.";
RL Arerugi 43:435-440(1994).
RN [3]
RP DISULFIDE BONDS; AND PARTIAL SEQUENCE.
RX MEDLINE=93283958; PubMed=8508052;
RA Nishiyama C., Yuuki T., Takai T., Okumura Y., Okudaira H.;
RT "determination of three disulfide bonds in a major house dust mite
RT allergen, Der f II.";
RL Int. Arch. Allergy Immunol. 101:159-166(1993).
RN [4]
RP PARTIAL SEQUENCE OF 18-52.
RX MEDLINE=89278484; PubMed=2732406;
RA Heymann P.W., Chapman M.D., Aalberse R.C., Fox J.W.,
RA Platts-Mills T.A.;
RT "Antigenic and structural analysis of group II allergens (Der f II
RT and Der p II) from house dust mites (Dermatophagoides spp).";
RL J. Allergy Clin. Immunol. 83:1055-1067(1989).

```

```

RN  STRUCTURE BY NMR.
RX  MEDLINE=98079068; PubMed=9417088;
RA  Ichikawa S., Hatanaka H., Yuuki T., Iwamoto N., Kojima S.,
RA  Nishiyama C., Ogura K., Okumura Y., Inagaki F.;
RT  "Solution structure of Der f 2, the major mite allergen for atopic
RT  diseases.";
RL  J. Biol. Chem. 273:356-360(1998).
CC  -!- SUBCELLULAR LOCATION: Secreted.
CC  -!- MISCELLANEOUS: THE SEQUENCE SHOWN HERE IS FROM CLONE 2. THE
CC  N-TERMINAL SEQUENCE (AA 1-8) FROM CLONE 1 AND 11 ARE NOT YET
CC  KNOWN.
CC  -!- SIMILARITY: BELONGS TO THE NPC2 FAMILY.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  EMBL: D10447; BAA01239.1; -
CC  DR  EMBL: D10448; BAA01240.1; -
CC  DR  EMBL: D10449; BAA01241.1; -
CC  DR  EMBL: S70378; AAB30829.1; -
CC  DR  PIR: PS0417; PS0417.
CC  DR  PDB: 1AHM; 08-APR-98.
CC  DR  PDB: 1AHM; 08-APR-98.
CC  DR  InterPro: IPR003172; El_DerP2_DerF2.
CC  DR  Pfam: PF02221; El_DerP2_DerF2; 1.
KW  Allergen; Signal; Polymorphism; 3D-structure.
FT  SIGNAL 1 17
FT  CHAIN 18 146 MITE GROUP 2 ALLERGEN DER F 2.
FT  DISULFID 25 136
FT  DISULFID 38 44
FT  DISULFID 90 95
FT  VARIANT 93 103 M -> V (IN CLONE 1).
FT  VARIANT 105 105 I -> A (IN CLONE 11).
FT  VARIANT 128 128 I -> V (IN CLONE 11).
FT  VARIANT 142 142 G -> A (IN CLONE 11).
FT  CONFLICT 5 8 ILCL -> GTMV (IN REF. 2).
SQ  SEQUENCE 146 AA; 15802 MW; FA118206CD88534A CRC64;

Query Match 12.2%; Score 97; DB 1; Length 146;
Best Local Similarity 88.2%; Pred. No. 0.005;
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 52 GCHGSEPCIHRGKPF 68
DB 37 GCHGSDPCIHRGKPF 53
|||||:|||||:|||||:
|||||:|||||:|||||:

RESULT 4
BAR3_CHITE
ID BAR3_CHITE STANDARD; PRT: 1700 AA.
AC Q03376;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Balbani ring protein 3 precursor.
GN BR3.
OS Chironomus tentans (Midge).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Chironomidae; Chironomidae; Chironominae; Chironomus.
OX NCBI_TaxID=7153;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=90172404; PubMed=1689777;
RA Paulsson G., Lendahl U., Galli J., Ericsson C., Wieslander L.;
RT "The Balbani ring 3 gene in Chironomus tentans has a diverged

```

```

RT repetitive structure split by many introns.";
RL J. Mol. Biol. 211:331-349(1990).
CC -!- FUNCTION: USED BY THE LARVAE TO CONSTRUCT A SUPRAMOLECULAR
CC STRUCTURE, THE LARVAL TUBE. BALBIANI RING PROTEIN 3 COULD PLAY A
CC ROLE AS A TRANSPORT PROTEIN THAT BINDS TO OTHER PROTEINS
CC INTRACELLULARLY AND IN THE GLAND LUMEN IN ORDER TO PREVENT THESE
CC FROM FORMING WATER-INSOLUBLE FIBERS TOO EARLY.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: SALIVARY GLAND.
CC -!- DOMAIN: HAS 82 APPROXIMATE REPEATS OF CYS-X-CYS-X-CYS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: X52263; CAA36506.1; -
CC DR  PIR: S08167; S08167.
CC DR  HSSP: P15358; ISKZ.
CC DR  InterPro: IPR004153; CXXC_repeat.
CC DR  Pfam: PF03128; CXXC; 71.
KW Repeat; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 1700 BALBIANI RING PROTEIN 3.
SQ SEQUENCE 1700 AA; 186145 MW; 34202B28521B0815 CRC64;

Query Match 11.3%; Score 90; DB 1; Length 1700;
Best Local Similarity 22.8%; Pred. No. 0.25;
Matches 37; Conservative 10; Mismatches 49; Indels 66; Gaps 7;

QY 35 SCHGSEPCIHRGKPF 74
DB 1152 SCCKNP-----KPANGCTGVQEWNEKEKCECPKDKPKKQCPGGQDNNHQCQCGCPT 1205
|||||:|||||:|||||:
|||||:|||||:|||||:

QY 75 -----PCIHRGKPF 105
DB 1206 PAPTCSNNQKYSVSCGCGNPGKPGKNGPQNIWCDNTCRVCVCPKNMEKPADNCKTKWW 1265
|||||:|||||:|||||:
|||||:|||||:|||||:

QY 106 GSE--PCIHRGKPF 136
DB 1266 NDEMCQCVKPGCGPEGGCKGVKWNANTCCECPADKAKPAS 1307
|||||:|||||:|||||:
|||||:|||||:|||||:

RESULT 5
Z071_XENLA
ID Z071_XENLA STANDARD; PRT: 898 AA.
AC P18751;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Oocyte zinc finger protein XLCOP7.1 (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE OF 1-311 FROM N.A.
RX MEDLINE=89345612; PubMed=2503827;
RA Knoecheil W., Poeting A., Koester M., el Baradi T., Nietfeld W.,
RA Bouwmeester T., Pieler T.;
RT "Evolutionary conserved modules associated with zinc fingers in
RT Xenopus laevis.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:6097-6100(1989).
RN [2]
RP SEQUENCE OF 284-898 FROM N.A.
RX MEDLINE=90040698; PubMed=2509712;
RA Nietfeld W., El-Baradi T., Mentzel H., Pieler T., Koester M.,
RA Poeting A., Knoecheil W.;
RT "Second-order repeats in Xenopus laevis finger proteins.";

```

```
RL J. Mol. Biol. 208:639-659(1989).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M25866; AA50013.1; -.
DR PIR; A33282; A33282.
DR PIR; S06546; S06546.
DR HSSP; P08047; 1SP2.
DR InterPro: IPR000822; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 21.
DR SMART; SM00355; ZINC_FINGER_C2H2_1; 21.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 21.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 21.
KW Zinc-finger; Metal-binding; DNA-binding; Repeat.
FT ZN_FING 289 311
FT ZN_FING 342 364
FT ZN_FING 370 392
FT ZN_FING 398 420
FT ZN_FING 425 448
FT ZN_FING 426 448
FT ZN_FING 454 476
FT ZN_FING 481 504
FT ZN_FING 482 504
FT ZN_FING 510 532
FT ZN_FING 537 561
FT ZN_FING 538 561
FT ZN_FING 567 589
FT ZN_FING 595 617
FT ZN_FING 623 645
FT ZN_FING 651 673
FT ZN_FING 679 701
FT ZN_FING 707 729
FT ZN_FING 735 758
FT ZN_FING 764 786
FT ZN_FING 792 814
FT ZN_FING 820 842
FT ZN_FING 848 870
FT ZN_FING 876 898
FT NON_TER 898 898
SQ SEQUENCE 898 AA; 101167 MW; 795D806E5696B0B6 CRC64;

Query Match 11.1%; Score 88; DB 1; Length 898;
Best Local Similarity 24.3%; Pred. No. 0-21;
Matches 44; Conservative 13; Mismatches 60; Indels 64; Gaps 12;

QY 19 QYIKANSKFTIGITELSSCHSEPCI-----IHRG-KPF-----GGCHGSEPCIHR----- 63
Db 614 QMHTGKPFSCBECGCFASSSDLTFRHTHTGKPFSCBECGCKVSKKSLVHQRTH 673
QY 64 --GKPFSSCHGSEPCI-----IHRGKPFGGCHGSEPC-----IHRGK 99
Db 674 TGEKPF-SCSKCDKCFASSSELNTHRTHTGKAFSCBECGCKFTNRSQLSRQHMHTGE 732
QY 100 PFSSCHGSEPCI-----HR-GKPF-----GGCHGSEPC-----CIIHRG-KPF 135
Db 733 KPTSCPECECFVSSSLTAHQQAHRMVPFSCLECKCFNSRNSNFARQHMHTGKPF 792
QY 136 s 136
Db 793 s 793

RESULT 6
ID F92_MOUSE
AC Q62396; STANDARD; PRT; 488 AA.
DT 15-DEC-1998 (Rel. 37, Created)
```

```
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Zinc finger protein 92 (Zfp-92).
GN ZFP92.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA MEDLINE=96425694; PubMed=8828036;
RA Levin M.L., Chatterjee A., Pragliola A., Worley K.C., Wehnert M.,
RA Zhuchenko O., Smith R.F., Lee C.C., Herman G.E.;
RA "A comparative transcription map of the murine bare patches (Bpa) and
RT striated (Str) critical regions and human Xq28.";
RL Genome Res. 6:465-477(1996).
CC -!- FUNCTION: MAY FUNCTION AS A TRANSCRIPTION FACTOR.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -!- SIMILARITY: CONTAINS 1 KRAB DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U47104; AAC52629.1; -.
DR HSSP; P25490; IUBD.
DR MGD; MGI:108094; Zip92.
DR InterPro: IPR001909; KRAB.
DR InterPro; IPR000822; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 9.
DR Pfam; PF01352; KRAB; 1.
DR PRINTS; PR00048; ZINCFINGER.
DR PRODOM; PD000003; Znf_C2H2; 4.
DR SMART; SM00349; KRAB; 1.
DR SMART; SM00355; Znf_C2H2; 9.
DR PROSITE; PS00805; KRAB; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 8.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 9.
DR Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
KW Nuclear protein; Repeat.
FT DOMAIN 14 85
FT DOMAIN 141 432
FT ZN_FING 141 163
FT ZN_FING 169 191
FT ZN_FING 197 219
FT ZN_FING 225 247
FT ZN_FING 253 275
FT ZN_FING 281 303
FT ZN_FING 337 359
FT ZN_FING 410 432
SQ SEQUENCE 488 AA; 55961 MW; 6A649E30F2043699 CRC64;

Query Match 10.8%; Score 86; DB 1; Length 488;
Best Local Similarity 33.6%; Pred. No. 0-19;
Matches 49; Conservative 10; Mismatches 39; Indels 48; Gaps 15;

QY 14 TELGGQYIKANSKFTIGITELSSCH-GSEPCIHR-GKPFGG----CHGSEPCIHRG-KP 66
Db 256 TEGC----KAFSSRNLIHQRIHSGQPKYICKGKAFKGVSVIHHQ---LIHRGDKP 308
QY 67 FSSCHGSEPCIHRCKPFGGCHG-SEPCIHRG-KPF--SSCHGSEPCIHRGKPFGG-- 120
Db 309 F-TCH-----EYKAFRGLSGLSQHQRVHRGKPYECSEC-----GRAFGRA 350
QY 121 -----CHGSEPCIHR--RGKPFGR 137
```

Db 351 NLFKHQVHGGVVR-LQHRTRGKGFQR 375

RESULT 7

SP87\_DICDI STANDARD; PRT; 677 AA.

AC P54643;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE Spore coat protein SP87 precursor (PL3 protein).

GN PSPD.

OS Dictyostelium discoideum (Slime mold).

OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.

OX NCBI\_TaxID=44689;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=AX3;

RX MEDLINE=9422358; PubMed=8174787;

RA Yoder B.K., Mao J., Erdos G.W., Blumberg D.D.:

RT "Identification of a new spore coat protein gene in the cellular slime mold Dictyostelium discoideum.";

RL Dev. Biol. 163:49-65(1994).

CC -!- FUNCTION: MAY CONTRIBUTE TO THE STRUCTURE OF THE COAT AT THE INTERFACE BETWEEN THE MIDDLE, CELLULOSIC LAYER AND THE OUTER, ELECTRON-DENSE, PROTEINACEOUS LAYER.

CC -!- SUBCELLULAR LOCATION: ACCUMULATES SPECIFICALLY IN REGULATED SECRETORY VESICLES OF PRESPORE CELLS (PRESPORE VESICLES). THE PROTEIN LATER ACCUMULATES EXTRACELLULARLY IN THE SPORE COAT.

CC -!- INDUCTION: BY C-AMP.

CC -!- PTM: DISULFIDE BONDING IS IMPORTANT FOR ASSOCIATING SP87 WITH THE COAT.

CC -!- PTM: PHOSPHORYLATED ON SERINE RESIDUES.

CC -!- SIMILARITY: CONTAINS 12 PRESPORE MOTIFS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; U25144; AAA73515.1; -.

DR Dictyob; DD02054; pspd.

DR InterPro; IPR003645; FOLN.

DR SMART; SM00274; FOLN: 5.

KW Glycoprotein; Repeat; Sporulation; Signal.

FT SIGNAL 1 21

FT CHAIN 22 677 SPORE COAT PROTEIN SP87.

FT REPEAT 189 201 PRESPORE MOTIF 1.

FT REPEAT 232 244 PRESPORE MOTIF 2.

FT REPEAT 276 288 PRESPORE MOTIF 3.

FT REPEAT 316 328 PRESPORE MOTIF 4.

FT REPEAT 356 368 PRESPORE MOTIF 5.

FT REPEAT 390 402 PRESPORE MOTIF 6.

FT REPEAT 418 430 PRESPORE MOTIF 7.

FT REPEAT 450 462 PRESPORE MOTIF 8.

FT REPEAT 513 525 PRESPORE MOTIF 9.

FT REPEAT 543 555 PRESPORE MOTIF 10.

FT REPEAT 571 583 PRESPORE MOTIF 11.

FT REPEAT 601 613 PRESPORE MOTIF 12.

FT DOMAIN 148 192 5 X 9 AA REPEATS OF G-G-S-S-G-G-T-S.

FT REPEAT 148 156 1-1 (APPROXIMATE).

FT REPEAT 157 165 1-2.

FT REPEAT 166 174 1-3.

FT REPEAT 175 183 1-4 (APPROXIMATE).

FT REPEAT 184 192 1-5 (APPROXIMATE).

FT DOMAIN 268 307 3 X 40 AA APPROXIMATE REPEATS.

FT REPEAT 268 307 2-1.

FT REPEAT 308 347 2-2.

FT REPEAT 348 387 2-3.

FT DOMAIN 391 628 8 X 26 AA APPROXIMATE REPEATS.

FT REPEAT 391 417 3-1.

FT REPEAT 419 445 3-2.

FT REPEAT 451 477 3-3.

FT REPEAT 480 506 3-4.

FT REPEAT 514 540 3-5.

FT REPEAT 544 570 3-6.

FT REPEAT 572 598 3-7.

FT REPEAT 602 628 3-8.

FT DOMAIN 141 185 GLY/SER/THR-RICH.

FT DOMAIN 256 271 GLY/SER/THR-RICH.

FT DOMAIN 342 351 GLY/SER/THR-RICH.

FT DOMAIN 174 183 POLY-SER.

FT DOMAIN 445 448 POLY-PRO.

FT CARBOHYD 72 72 N-LINKED (GLCNAC... ) (POTENTIAL).

SQ SEQUENCE 677 AA; 71804 MW; EF6E29CF57E78D5 CRC64;

Query Match 10.8%; Score 86; DB 1; Length 677;

Best Local Similarity 29.2%; Pred. No. 0.25;

Matches 38; Conservative 10; Mismatches 46; Indels 36; Gaps 12;

QY 37 HGSEPCIIHRGKPFPGG---HGSEPCIIHRGKPFSSCHGSEPCIIHRGKPFPGC-----HG 89

Db 475 HCDEVCDLDCGRGF-ECKIRHDSKCCVRSERPHPPQH--EKC-NKRCPPGHECKVDQHG 530

QY 90 SEPCII-HRGKPFSS-----CH-----GSEPCI-THRGK-----PFGCC-----HGSE 125

Db 531 KECVVAHRPPPKCSLRCPPRHCVNHFGECCVKVHHDKCSLRCPPGHECKVDQHGKE 590

QY 126 PCII-HRGK 134

Db 591 CCVVAHRPPP 600

RESULT 8

CORB\_MOUSE

ID CORB\_MOUSE STANDARD; PRT; 153 AA.

AC Q62267;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Cornifin B (Small proline-rich protein 1B) (SPR1B) (SPR1 B).

GN SPR1B.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CD-1;

RX MEDLINE=96179743; PubMed=8601731;

RA Kartasova T., Darwiche N., Kohno Y., Koizumi H., Osada S.-I.,

RA Huh N.-H., Lichti U., Steinert P.M., Kuroki T.;

RT "Sequence and expression patterns of mouse SPR1: correlation of expression with epithelial function.";

RL J. Invest. Dermatol. 106:294-304(1996).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=129/SvJ;

RA Reddy P.M.S., Wu R.;

RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: CROSS-LINKED ENVELOPE PROTEIN OF KERATINOCYTES. IT IS A KERATINOCYTE PROTEIN THAT FIRST APPEARS IN THE CELL CYTOSOL, BUT ULTIMATELY BECOMES CROSS-LINKED TO MEMBRANE PROTEINS BY TRANSGLUTAMINASE. ALL THAT RESULTS IN THE FORMATION OF AN INSOLUBLE ENVELOPE BENEATH THE PLASMA MEMBRANE.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC -!- TISSUE SPECIFICITY: EXPRESSED IN FETAL PTERIDERM, HAIR FOLLICLES AND IN THE THICKENED EPIDERMIS OF THE LIP AND FOOTPAD. ALSO PRESENT IN THE EPITHELIA OF VARIOUS TISSUES SUCH AS THE PENIS, VAGINA, FORESTOMACH, TONGUE AND ESOPHAGUS.

CC -!- DEVELOPMENTAL STAGE: FIRST DETECTED IN FETAL SKIN AROUND DAY 16 AND EXPRESSION CONTINUES THROUGHOUT NEWBORN AND ADULT STAGES.

CC -!- SIMILARITY: BELONGS TO THE CORNIFIN (SPRR) FAMILY.





[illegible]



```

FT CARBOHYD 233 233 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 421 421 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 431 431 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 466 466 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 648 648 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 769 AA: 85631 MW; F7E3994F92B12A65 CRC64;

Query Match 10.0%; Score 79; DB 1; Length 769;
Best Local Similarity 25.8%; Pred. No. 1.4;
Matches 42; Conservative 14; Mismatches 55; Indels 52; Gaps 11;

Qy 9 KFIGITELGGVYKAN-----SKFIGITELSS-----CH-----GSEP 41
Db 451 KPIGNETAKIHHNCSQCQEDNRPKQKCVDEFTLDSKFCQCDNKCHDEDQSSSES 510
Qy 42 CLIHKGKPGGCHGSEPCIIHGRKPFSSCHGSEPCIIHGRKPFSSCHGSEPCIIHGRK 99
Db 511 CKSHKQDPV--CSGRGVCV--GK--CSCHK-----IKLGKYGKCYCKDDPCPYHGG-- 558
Qy 100 PFSSCHGSEPCIIHGRKPFSSCHGCHG-----SEPCIIHGRK 133
Db 559 --NLCAGHGECEAGRCQCFSGWEGDRQCPSAAAHQHCVNKSGQ 599

RESULT 14
TETX_CLOTE
ID TETX_CLOTE STANDARD; PRT; 1314 AA.
AC P04958;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Tetanus toxin precursor (EC 3.4.24.68) (Tentoxylisin).
OS Clostridium tetani.
OG Plasmid.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87053814; PubMed=3536478;
RA Eisel U., Jarausch W., Goretzki K., Henschen A., Engels J.,
RA Weller U., Hudel M., Habermann E., Niemann H.;
RT "Tetanus toxin: primary structure, expression in E. coli, and
RT homology with botulinum toxins.";
RL EMBO J. 5:2495-2502(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=CN3911;
RX MEDLINE=87040747; PubMed=3774547;
RA Fairweather N.F., Lyness V.A.;
RA "The complete nucleotide sequence of tetanus toxin.";
RL Nucleic Acids Res. 14:7809-7812(1986).
RN [3]
RP SEQUENCE OF 742-1314 FROM N.A.
RX MEDLINE=86085672; PubMed=3510187;
RA Fairweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.;
RT "Cloning, nucleotide sequencing, and expression of tetanus toxin
RT fragment C in Escherichia coli.";
RL J. Bacteriol. 165:21-27(1986).
RN [4]
RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=90201034; PubMed=2108021;
RA Kriegstein K., Henschen A., Weller U., Habermann E.;
RT "Arrangement of disulfide bridges and positions of sulfhydryl groups
RT in tetanus toxin.";
RL Eur. J. Biochem. 188:39-45(1990).
RN [5]
RP PARTIAL SEQUENCE.
RX MEDLINE=92037649; PubMed=1935979;
RA Kriegstein K.G., Henschen A.H., Weller U., Habermann E.;
RT "Limited proteolysis of tetanus toxin. Relation to activity and

```

```

RT identification of cleavage sites.";
RL Eur. J. Biochem. 202:41-51(1991).
RN [6]
RP IDENTIFICATION AS ZINC-PROTEASE.
RX MEDLINE=93010948; PubMed=1396558;
RA Schiavo G., Poulain B., Rossetto O., Benfenati F., Tauc L.,
RA Montecucco C.;
RT "Tetanus toxin is a zinc protein and its inhibition of
RT neurotransmitter release and protease activity depend on zinc.";
RL EMBO J. 11:3577-3583(1992).
RN [7]
RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=93063293; PubMed=1331807;
RA Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,
RA Dasgupta B.R., Montecucco C.;
RT "Tetanus and botulinum-B neurotoxins block neurotransmitter release
RT by proteolytic cleavage of synaptobrevin.";
RL Nature 359:832-835(1992).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 874-1314.
RX MEDLINE=97475217; PubMed=9334741;
RA Umland T.C., Wingert L.M., Swaminathan S., Furey W.F., Schmidt J.J.,
RA Sax M.;
RT "Structure of the receptor binding fragment HC of tetanus
RT neurotoxin.";
RL Nat. Struct. Biol. 4:788-792(1997).
CC -I- FUNCTION: TETANUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 76-GLN-1-PHE-77
CC BOND OF SYNAPTOSOM-2.
CC -I- CATALYTIC ACTIVITY: HYDROLYSIS OF 76-GLN-1-PHE-77 BOND IN
CC SYNAPTOSOM-2.
CC -I- SUBUNIT: THE PRECURSOR POLYPEPTIDE IS SUBSEQUENTLY CLEAVED TO
CC YIELD SUBCHAINS L AND H. THESE REMAIN LINKED BY A DISULFIDE BRIDGE
CC AND ARE NON-TOXIC AFTER SEPARATION.
CC -I- MISCELLANEOUS: THE C-TERMINAL OF THE HEAVY CHAIN BINDS TO
CC GANGLIOSIDE RECEPTORS.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X04436; CAA28033.1; -
DR EMBL; M12739; AAA23282.1; -
DR EMBL; X06214; CAA29564.1; -
DR PIR; A25689; BTCLTN.
DR PDB; 1AF9; 29-APR-98.
DR PDB; 1A8D; 14-OCT-98.
DR MEROPS; M27_001; -
DR InterPro; IPR000395; Bontoxilysin.
DR InterPro; IPR000130; Zn_Mtpeptidse.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOXILYSIN.
DR ProDom; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc; Plasmid;
KW 3D-structure.
FT INIT_MET 0
FT CHAIN 1 456 TETANUS TOXIN LIGHT CHAIN.
FT CHAIN 457 1314 TETANUS TOXIN HEAVY CHAIN.
FT METAL 232 232 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 233 233 BY SIMILARITY.
FT METAL 236 236 ZINC (CATALYTIC) (BY SIMILARITY).
FT TRANSMEM 226 246 POTENTIAL.
FT TRANSMEM 669 689 POTENTIAL.

```

```

use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; U28136; AAA68953.1; -
EMBL; U18353; AAC47458.1; -
EMBL; M14778; AAC28644.1; -
EMBL; M13568; AAA28645.1; -
PIR; A26378; A26378.
HSP; P06213; IIRK.
FlyBase; FBgn0013984; InR.
InterPro; IPR000494; EGFR_L_domain.
InterPro; IPR000719; Euk_pkinase.
InterPro; IPR003961; FN_III.
InterPro; IPR002174; Furin-like.
InterPro; IPR002011; RTKinaseII.
InterPro; IPR001245; Tyr_pkinase.
Pfam; PF00069; pkinase; 1.
Pfam; PF00757; Furin-like; 1.
Pfam; PF01030; Recep_L_domain; 2.
PRINTS; PR00109; TYRKINASE.
ProDom; PD000001; Euk_pkinase; 1.
SMART; SM00060; FN3; 1.
SMART; SM00261; FU; 1.
SMART; SM00219; TyrKc; 1.
PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
Kw Transferrase; Tyrosine-protein kinase; Receptor; Transmembrane;
Kw Glycoprotein; ATP-binding; Phosphorylation; Signal.
FT SIGNAL 1 ?
FT CHAIN ? 2146 INSULIN-LIKE RECEPTOR.
FT CHAIN 651 1077 ALPHA-SUBUNIT.
FT CHAIN 1082 2146 BETA-SUBUNIT.
FT DOMAIN 1082 1307 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1308 1332 POTENTIAL.
FT DOMAIN 1333 2146 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 1368 1655 PROTEIN KINASE.
FT NP_BIND 1374 1382 ATP (BY SIMILARITY).
FT BINDING 1402 1402 ATP (BY SIMILARITY).
FT BINDING 1516 1516 BY SIMILARITY.
FT ACT_SITE 1546 1546 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD 74 74 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 203 203 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 265 265 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 376 376 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 468 468 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 509 509 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 561 561 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 569 569 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 751 751 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 810 810 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 824 824 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 839 839 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 898 898 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 943 943 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1049 1049 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1143 1143 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1214 1214 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1262 1262 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 29 29 T -> ATTAK (IN REF. 2).
FT CONFLICT 89 89 D -> V (IN REF. 2).
FT CONFLICT 164 164 C -> F (IN REF. 2).
FT CONFLICT 470 470 S -> T (IN REF. 2).
FT CONFLICT 485 485 S -> R (IN REF. 2).
FT CONFLICT 653 653 DSERRARE -> PPPPPPL (IN REF. 3).
FT CONFLICT 679 681 RES -> GER (IN REF. 3).
FT CONFLICT 723 727 ISGDP -> LAAI (IN REF. 3).
FT CONFLICT 793 793 E -> V (IN REF. 2 AND 3).

```

```
FT CONFLICT 822 822 S -> M (IN REF. 2 AND 3).
FT CONFLICT 864 877 TOLKAVTIHAMIAG -> NSTKSSDDPCDDRW (IN
FT REF. 2).
FT CONFLICT 932 932 V -> VTEV (IN REF. 2).
FT CONFLICT 951 954 NLMA -> KPYGV (IN REF. 2).
FT CONFLICT 1157 1157 MISSING (IN REF. 3).
FT CONFLICT 1183 1188 LCSDYD -> SAATIH (IN REF. 3).
FT CONFLICT 1220 1229 ATFSLGRHOL -> VYRVTPPV (IN REF. 2).
FT CONFLICT 1260 1279 DFQTAGYLLKNEGLYSFR -> RLQPCDRLFNKAQGRPL
FT OLQ (IN REF. 3).
FT CONFLICT 1297 1300 IKVE -> LIOQ (IN REF. 4).
FT CONFLICT 1454 1455 GD -> VE (IN REF. 3).
FT CONFLICT 1466 1481 EERDEAMMYLNIRIGV -> PRSGMRPDDVSLIAMM (IN
FT REF. 3).
FT CONFLICT 1496 1496 M -> V (IN REF. 3).
FT CONFLICT 1516 1517 DL -> PF (IN REF. 4).
FT CONFLICT 1566 1575 RDGYSSASD -> QAWCLLVPVT (IN REF. 4).
FT CONFLICT 1588 1595 TLAQPYQ -> ILSLWRSP (IN REF. 4).
FT CONFLICT 1679 1679 N -> H (IN REF. 2 AND 3).
FT CONFLICT 1703 1703 T -> S (IN REF. 3).
FT CONFLICT 1712 1712 Q -> E (IN REF. 2).
FT CONFLICT 1849 1851 FTT -> SAA (IN REF. 2).
FT CONFLICT 1884 1884 Y -> F (IN REF. 2).
FT CONFLICT 2092 2146 TKRENLRIPTPESEROEPTDPTCLKRTGCDRVRROGLHH
FT PMDSSEGRNRREL -> SOTRKSPTNPNSGIGATGAGNRS
FT NLLKENW LRPASTPREP PPNGFIGREA (IN
FT REF. 2).
SQ SEQUENCE 2146 AA; 240488 MW; CC83B42654768E9D CRC64;
```

```
Query Match 9.8%; Score 78; DB 1; Length 2146;
Best Local Similarity 23.3%; Pred. No. 4.6;
Matches 35; Conservative 13; Mismatches 38; Indels 64; Gaps 9;
```

```
QY 30 ITELSSCHGSEPCIIHRGKPPGCG-----HGSEPCIIHRGKPF-----SS 69
Db 540 IDEHTCC--SQDCL-----GGCVIDKNGNESCISCRNVSNFNNICMDCSPKGYVQFDSR 590
QY 70 CHGSEPCI-----IHRGKPGGCHGSEPCIIHRGKPFSSCHGS---BPC-----110
Db 591 CVTANECITLTKFETNSVSGIPYNG-----QCITHCPTGYQKSENKRMCEPCPGGKCDK 645
QY 111 -----IHRGKPPGCG---HGSEPCII 129
Db 646 ECSSGLIDSLERAREFHGCHIIITGTEPLTI 675
```

Search completed: December 4, 2002, 12:53:29  
Job time : 16.8417 secs







```
Matches 15; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 52 GCHGSEPCIIHRGKPPS 68
Db 61 GCHGSDPCIIHRGKPPT 77

RESULT 6
QBR4Y4
ID QBR4Y4 PRELIMINARY; PRT; 2571 AA.
AC QBR4Y4;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Stabilin-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/C; TISSUE=LIVER;
RX MEDLINE=21818465; PubMed=11829752;
RA Politz O., Gratchev A., McCourt P.A.G., Schledzewski K., Guillot P.,
RA Johansson S., Svineng G., Franke P., Kannicht C., Kzhyskowska J.,
RA Longati P., Velten F.W., Johansson S., Goerdts S.;
RT "Stabilin-1 and -2 constitute a novel family of fasciclin-like
RT hyaluronan receptor homologues.";
RL Biochem. J. 362:155-164(2002).
DR EMBL; AF290914; AAL91671.2; -.
SQ SEQUENCE 2571 AA; 276254 MW; 77D00B943DC47718 CRC64;

Query Match 12.1%; Score 96; DB 11; Length 2571;
Best Local Similarity 24.8%; Pred. No. 0.062;
Matches 35; Conservative 12; Mismatches 52; Indels 42; Gaps 8;
QY 12 GITELG-----GOYIKANSFIGITELSS-CHGSEPCIIHRGKPGGCHGSEPCIIHR 63
Db 800 GVCQGTGAPGQGRFESMGNCSTGUAQPCSDAHCVIQGVARVCYCHDG-----FE 854
QY 64 GRPFSSCHGSEPCIIHRGKP-FGGCHGSEPCIIHRGKPFSSCHGSEPCIIHRG----- 115
Db 855 GNGF-SCRNSNFC-----SRPDGGGSENAECV-----PGDLGTHHCCHKGWSGDGRI 902
QY 116 -----KPGGGCHGSPC 127
Db 903 CVAIDECGLDTRGGCHADALC 923

RESULT 7
Q9VR08
ID Q9VR08 PRELIMINARY; PRT; 3680 AA.
AC Q9VR08;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE CG15637 protein.
GN DP OR CG15637.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OX Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S.D., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
```

```
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballaw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Bottlier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclebb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J., Yao Q.A.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003576; AAF51000.1; -.
DR HSSP; P35555; IEMN.
DR FlyBase; FBgn0000488; dp.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR002603; ET.
DR InterPro; IPR003645; FOLN.
DR InterPro; IPR002899; WR1/EB.
DR Pfam; PF00008; EGF; 25.
DR ProDom; PD003573; ET; 1.
DR SMART; SM00179; EGF_CA; 16.
DR SMART; SM00001; EGF_like; 35.
DR SMART; SM00274; FOLN; 5.
DR SMART; SM00289; WR1; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 18.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 18.
DR PROSITE; PS01187; EGF_CA; 17.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 3680 AA; 388637 MW; E5E972E1A3479EFF CRC64;

Query Match 11.9%; Score 94.5; DB 5; Length 3680;
Best Local Similarity 23.8%; Pred. No. 0.14;
Matches 30; Conservative 15; Mismatches 52; Indels 29; Gaps 6;
QY 25 SKFTIGITELSSCHGSEPCIIHRGKPGGCHGSEPCIIHRGKPFSS----CHGSEPCIIHR 80
Db 707 SKTVGCS DANPACGTGTCV-----QDSYTGNSVCICRQGYERNSENGCQDVCDSVOR 760
QY 81 GKPFPGGCH-----GSEPCI-----HRGKPFSSCH--GSEPC-----IHRGKPFGGC 121
Db 761 GKPFAGLNA LCKNLPGSYECRCPOGHNGNPFIMCEICNTPECCQCSPKLVGNSCVLSGC 820
QY 122 HGSEPC 127
Db 821 SSGQAC 826

RESULT 8
Q8R5B3
```

ID	Q8R5B3	PRELIMINARY;	PRT;	502 AA.
AC	Q8R5B3			
DT	01-JUN-2002 (TReMBLrel. 21, Created)			
DT	01-JUN-2002 (TReMBLrel. 21, Last sequence update)			
DT	01-JUN-2002 (TReMBLrel. 21, Last annotation update)			
DE	Hypothetical 58.5 kDa protein.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Strausberg R.			
RL	Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; BC023090; AAH23090.1; -			
KW	Hypothetical protein.			
SQ	SEQUENCE 502 AA; 58468 MW; 0995426E8A1D5BC2 CRC64;			
Query Match 11.7%; Score 92.5; DB 11; Length 502;				
Best Local Similarity 30.2%; Pred. No. 0.026;				
Matches 45; Conservative 11; Mismatches 38; Indels 55; Gaps 15;				
QY	38 GSEPCI-IHRGKPF---GGCHGSEPCI-IHRG-KPFS-----SCH-----GS 73			
Db	243 GKPIVCTHCGKAFIDRRTHNHER--THTGVKPFACKCGKAFRLRSCOLLIIHERIHTGE 300			
QY	74 EPCII-HRGKPF---GGCH-----GSEPCI-IHRG-KPFS-----SCHGSEPCI 111			
Db	301 RFVCKHCGKATYTSACYYHERIHTGKPCVKCKGKAFKCSAYLIHHERSHSGKPYV 360			
QY	112 I-HRGKPF---GGCHGSEPCI-IHRG-KPF 135			
Db	361 CRKCGKAFAYATGCHKHER--IHTGKPY 387			
RESULT 9				
Q8R2M7	PRELIMINARY;	PRT;	528 AA.	
ID	Q8R2M7			
AC	Q8R2M7			
DT	01-JUN-2002 (TReMBLrel. 21, Created)			
DT	01-JUN-2002 (TReMBLrel. 21, Last sequence update)			
DE	Similar to zinc finger protein sil-6.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Strausberg R.			
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; BC027407; AAH27407.1; -			
SQ	SEQUENCE 528 AA; 61431 MW; 8A586471C4D2C88A CRC64;			
Query Match 11.7%; Score 92.5; DB 11; Length 528;				
Best Local Similarity 30.2%; Pred. No. 0.028; 38; Indels 55; Gaps 15;				
Matches 45; Conservative 11; Mismatches 38; Indels 55; Gaps 15;				
QY	38 GSEPCI-IHRGKPF---GGCHGSEPCI-IHRG-KPFS-----SCH-----GS 73			
Db	243 GKPIVCTHCGKAFIDRRTHNHER--THTGVKPFACKCGKAFRLRSCOLLIIHERIHTGE 300			
QY	74 EPCII-HRGKPF---GGCH-----GSEPCI-IHRG-KPFS-----SCHGSEPCI 111			
Db	301 RFVCKHCGKATYTSACYYHERIHTGKPCVKCKGKAFKCSAYLIHHERSHSGKPYV 360			
QY	112 I-HRGKPF---GGCHGSEPCI-IHRG-KPF 135			
Db	361 CRKCGKAFAYATGCHKHER--IHTGKPY 387			
RESULT 10				
Q8TLW0	PRELIMINARY;	PRT;	108 AA.	
ID	Q8TLW0			
AC	Q8TLW0			
DT	01-JUN-2002 (TReMBLrel. 21, Created)			
DT	01-JUN-2002 (TReMBLrel. 21, Last sequence update)			
DT	01-JUN-2002 (TReMBLrel. 21, Last annotation update)			
DE	Predicted protein.			
GN	MA2913.			
OS	Methanosarcina acetivorans.			
OC	Archaea; Euryarchaeota; Methanococci; Methanosarcinales;			
OC	Methanosarcinaceae; Methanosarcina.			
OX	NCBI_TaxID=2214;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	STRAIN=C2A / ATCC 35395 / DSM 2834;			
RC	MEDLINE=21929760; PubMed=11932238;			
RX	Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,			
RA	Fitzhugh W., Calvo S., Engels R., Smirnov S., Athoor D., Brown A.,			
RA	Allen N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,			
RA	Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,			
RA	Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,			
RA	Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,			
RA	Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,			
RA	Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,			
RA	Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,			
RA	Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,			
RA	Metcalf W.W., Birten B.;			
RT	"The genome of Methanosarcina acetivorans reveals extensive metabolic			
RT	and physiological diversity."			
RL	Genome Res. 12:532-542(2002).			
DR	EMBL; AF010993; AAM06289.1; -			
KW	Complete proteome.			
SQ	SEQUENCE 108 AA; 12530 MW; A91E25423B9475B4 CRC64;			
Query Match 11.5%; Score 91.5; DB 17; Length 108;				
Best Local Similarity 18.9%; Pred. No. 0.0064;				
Matches 14; Conservative 29; Mismatches 30; Indels 1; Gaps 1;				
QY	65 KPSSCHGSEPCI-IHRGKPGGCHGSEPCI-IHRGKPPSSCHGSEPCI-IHRGKPGGCHGS 124			
Db	2 QPYSKNRTQNTATVLMQPYSKNRTQNTATVLMQPYSKNRTQNTAAVLMQPYSKNRT 61			
QY	125 E-PCIIHRGKPFSSR 137			
Db	62 QNATILHSSEHFA 75			
RESULT 11				
Q9VY49	PRELIMINARY;	PRT;	341 AA.	
ID	Q9VY49			
AC	Q9VY49;			
DT	01-MAY-2000 (TReMBLrel. 13, Created)			
DT	01-MAY-2000 (TReMBLrel. 13, Last sequence update)			
DT	01-JUN-2002 (TReMBLrel. 21, Last annotation update)			
DE	CG11674 protein.			
GN	CG11674			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BERKELEY;			
RX	MEDLINE=20196006; PubMed=10731132;			
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
RA	Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,			
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,			

ID	Q8TLW0	PRELIMINARY;	PRT;	108 AA.
AC	Q8TLW0			
DT	01-JUN-2002 (TReMBLrel. 21, Created)			
DT	01-JUN-2002 (TReMBLrel. 21, Last sequence update)			
DT	01-JUN-2002 (TReMBLrel. 21, Last annotation update)			
DE	Predicted protein.			
GN	MA2913.			
OS	Methanosarcina acetivorans.			
OC	Archaea; Euryarchaeota; Methanococci; Methanosarcinales;			
OC	Methanosarcinaceae; Methanosarcina.			
OX	NCBI_TaxID=2214;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	STRAIN=C2A / ATCC 35395 / DSM 2834;			
RC	MEDLINE=21929760; PubMed=11932238;			
RX	Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,			
RA	Fitzhugh W., Calvo S., Engels R., Smirnov S., Athoor D., Brown A.,			
RA	Allen N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,			
RA	Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,			
RA	Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,			
RA	Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,			
RA	Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,			
RA	Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,			
RA	Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,			
RA	Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,			
RA	Metcalf W.W., Birten B.;			
RT	"The genome of Methanosarcina acetivorans reveals extensive metabolic			
RT	and physiological diversity."			
RL	Genome Res. 12:532-542(2002).			
DR	EMBL; AF010993; AAM06289.1; -			
KW	Complete proteome.			
SQ	SEQUENCE 108 AA; 12530 MW; A91E25423B9475B4 CRC64;			
Query Match 11.5%; Score 91.5; DB 17; Length 108;				
Best Local Similarity 18.9%; Pred. No. 0.0064;				
Matches 14; Conservative 29; Mismatches 30; Indels 1; Gaps 1;				
QY	65 KPSSCHGSEPCI-IHRGKPGGCHGSEPCI-IHRGKPPSSCHGSEPCI-IHRGKPGGCHGS 124			
Db	2 QPYSKNRTQNTATVLMQPYSKNRTQNTATVLMQPYSKNRTQNTAAVLMQPYSKNRT 61			
QY	125 E-PCIIHRGKPFSSR 137			
Db	62 QNATILHSSEHFA 75			
RESULT 11				
Q9VY49	PRELIMINARY;	PRT;	341 AA.	
ID	Q9VY49			
AC	Q9VY49;			
DT	01-MAY-2000 (TReMBLrel. 13, Created)			
DT	01-MAY-2000 (TReMBLrel. 13, Last sequence update)			
DT	01-JUN-2002 (TReMBLrel. 21, Last annotation update)			
DE	CG11674 protein.			
GN	CG11674			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BERKELEY;			
RX	MEDLINE=20196006; PubMed=10731132;			
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
RA	Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,			
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,			

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunko B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster K., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jajali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li X., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.:  
RA "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
DR EMBL: AE003495; AAF48356.1; -  
DR FlyBase: FBgn0030551; Cg11674.  
DR InterPro: IPR001304; Lectin.C.  
DR InterPro: IPR002899; WRI/EB.  
DR Pfam: PF01683; EB: 1.  
DR PROSITE: PS00615; C-TYPE LECTIN 1; UNKNOWN 1.  
SQ SEQUENCE 341 AA; 37711 MW; 70CDAAB4053B7B24 CRC64;

Query Match 11.5%; Score 91; DB 5; Length 341;  
Best Local Similarity 25.9%; Pred. No. 0.025;  
Matches 36; Conservative 12; Mismatches 45; Indels 46; Gaps 10;

QY 31 TELSSCHGSE-----PCIIHRGKPGGCHGSEPCIIHRGKPFSSCHGSEPCI- 77  
DB 86 TRNQCHCEGHVSSDRRRRCPLPAVVPVG---GSCFFQOOC---QRADRFSSCIGNQ-CLC 139  
QY 78 -----IHRG-----KPGGCHGSEPCIIHR-----GKPFSSCHGSEPCIIHR 114  
DB 140 LNQFPEHRCILSVLOSSCLEBDCGSC-GASICLTKTTCGCKNFVHNHNTKCI--K 196  
QY 115 GKPFGG-CHGSEPCIIHRG 132  
DB 197 GSAVGDTCHEHSPCKNLG 215

RESULT 12  
P90606  
ID P90606 PRELIMINARY; PRT; 270 AA.  
AC P90606;  
DT 01-MAY-1997 (Tremblrel. 03, Created)  
DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)  
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
DE Nucleic acid binding protein.  
OS Trypanosoma equiperdum.  
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.  
OX NCBI\_TaxID=5694;  
RN [1]  
RP Strand A.D.; Eisen H.;  
RT "Trypanosome CCHC zinc finger protein.";  
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U87108; AAB47542.1; -  
DR InterPro: IPR001878; Znf\_CCHC.  
DR Pfam; PF00098; zf-CCHC; 7.

DR PRINTS; PR00939; C2HCZNFINGER.  
DR SMART; SM00343; Znf\_C2HC; 7.  
SQ SEQUENCE 270 AA; 28453 MW; 50D2C559FD9F7234 CRC64;  
  
Query Match 11.3%; Score 89.5; DB 5; Length 270;  
Best Local Similarity 33.9%; Pred. No. 0.029;  
Matches 38; Conservative 4; Mismatches 43; Indels 27; Gaps 9;  
  
QY 37 HGSEPCIIHRG-KPGGCHGSEPCIIHRGKPFSSCHGSEPCIIHRGKPFGGC--HGSEPC 93  
DB 53 HLSRDCPSNRGTAPMGGRACYNC-----GQP---GHFSRECPNMRGGPMGGAPMGGRAC 105  
QY 94 IHRGKPFSSC-----HGSEPCIIHRGKPFGGC--HGSEPCIIHRGKPF--FSR 137  
DB 106 Y-----NCVQPGHFSRECPNMRGGPMGGAPMGGRAC-YHCGQPGHFSR 148

RESULT 13  
Q8WXC7  
ID Q8WXC7 PRELIMINARY; PRT; 389 AA.  
AC Q8WXC7;  
DT 01-MAR-2002 (Tremblrel. 20, Created)  
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
DE Hypothetical 42.2 kDa protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Hu W., Tang L.-J., Shi Y.-W., Tian J.-Y., Jian Y.-S.;  
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF453950; AAL41025.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 389 AA; 42241 MW; 3D72EE48EF92E5F5 CRC64;

Query Match 11.1%; Score 88; DB 4; Length 389;  
Best Local Similarity 25.5%; Pred. No. 0.063;  
Matches 38; Conservative 10; Mismatches 41; Indels 60; Gaps 11;

QY 37 HGSEPCIIHRGKPGGCHGSEPCIIHRGKPFSSCHGSEPCIIH-----RGKP 83  
DB 43 HGSSPMALH---PRLFTGSPSTALH---PCPFTHGSSPMSLHRLFTHSPPSMPLHPRP 96  
QY 84 FGGCHGS-----EPCIIHRG-----KPF---SSCHGSEPCIIHRG-----KPF 119  
DB 97 F--VHASSPTALRCLFTHGSPSMPLHPRPFVHASSPTALRCLFTHGSPSMPLHPRPF- 153  
QY 120 GCHGSE-----PCIIHRG-----KPF 135  
DB 154 -LHASSPTALRCLFTHGSPSMPLHPRPF 181

RESULT 14  
Q9NY15  
ID Q9NY15 PRELIMINARY; PRT; 2570 AA.  
AC Q9NY15;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
DE Stabilin-1.  
GN STAB1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Politz O., Guillot P., Gratchev A., Schledzewski K., Birk R.,  
RA Hakly N., Tebbe B., Orfanos C.E., Goerdts S.;  
RT "Stabilin-1: an endothelial-macrophage member of the fasciclin domain  
containing protein family associated with angiogenesis.";

RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ275213; CAB61827.1; -  
DR HSSP; P98066; ITSG.  
DR InterPro; IPR000782; BIGH3\_fasciclin.  
DR InterPro; IPR001128; Cytochrome\_P450.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR002049; Laminin\_EGF.  
DR InterPro; IPR000538; Link.  
DR Pfam; PF00008; EGF; 18.  
DR Pfam; PF02469; Fasciclin; 4.  
DR Pfam; PF00193; Xlink; 1.  
DR PRINTS; PR00011; EGF\_LAMININ.  
DR ProDom; PD000918; Link; 1.  
DR SMART; SM00180; EGF\_Lam; 1.  
DR SMART; SM00001; EGF\_Like; 15.  
DR SMART; SM00445; LINK; 1.  
DR PROSITE; PS00086; CYTOCHROME\_P450; UNKNOWN\_2.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_7.  
DR PROSITE; PS01186; EGF\_2; 16.  
DR PROSITE; PS01248; LAMININ\_TYPE\_EGF; 2.  
DR PROSITE; PS01241; LINK; UNKNOWN\_1.  
KW EGF-like domain; Glycoprotein; Laminin EGF-like domain; Repeat.  
SQ SEQUENCE 2570 AA; 275345 MW; 3123FABD7C8E2BF8 CRC64;

Query Match 11.1%; Score 88; DB 4; Length 2570;

Best Local Similarity 28.6%; Pred. No. 0.49;

Matches 34; Conservative 7; Mismatches 40; Indels 38; Gaps 10;

QY 35 SCHGSEPCIIHRGKPGGCH-----GSEPCIIHRGKPFSS--CHGSEPCIIHRGK 82

Db 858 SCTPSPNC-SHPDR--GGCSENACVPSGLGTHHCTCHKGWSDGRVCVAIDCELDVG- 913

QY 83 PFGGCH-----GSEPCIIHRGKPFSS--CHGSEPCIIHRGKPGGCHGSEPC 127

Db 914 --GGCHTDALCSVPGQSRCTCKLG--FAGDGYQCSPIIDPC--RAGN--GGCHGLATC 964

RESULT 15

Q9JJRO

ID Q9JJRO PRELIMINARY; PRT; 488 AA.

AC Q9JJRO;

DT 01-OCT-2000 (Tremblrel. 15, Created)

DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)

DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)

DE Zinc finger protein 92.

GN ZFP92.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=129/SV;

RA Mallon A.M., Platzer M., Bates R., Gloeckner G., Botcherby M.,

RA Nordsiek G., Strivens M.A., Kioschis P., Dangel A., Cunningham D.,

RA Straw R., Weston P., Hunter C., Gilbert M., Fernando S., Goodall K.,

RA Kerry G., Greystrom J.S., Clark D., Goerdes M., Blechschmidt K.,

RA Rump A., Hinzmann B., Mundy C.R., Miller W., Poustka A., Herman G.E.,

RA Rhodes M., Denny P., Rosenthal A., Brown S.D.M.;

RT "Comparative genome sequence analysis of the Bpa/Str region in mouse

and man."

RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.

CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).

DR EMBL; AL049866; CAB88276.1; -

DR HSSP; P25490; 1UBD.

DR MGD; MGI:108094; Zfp92.

DR InterPro; IPR001909; KRAB.

DR InterPro; IPR000822; Znf\_C2H2.

DR Pfam; PF01352; KRAB; 1.

DR Pfam; PF00096; Zf-C2H2; 9.

DR PRINTS; PR00048; ZINC\_FINGER.

DR ProDom; PD000003; Znf\_C2H2; 4.

DR SMART; SM00349; KRAB; 1.

DR

DR

DR

DR

DR SMART; SM00355; Znf\_C2H2; 9.

DR PROSITE; PS00805; KRAB; 1.

DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 8.

DR PROSITE; PS00157; ZINC\_FINGER\_C2H2\_2; 9.

KW DNA-binding; Metal-binding; Nuclear protein; Zinc-finger.

SQ SEQUENCE 488 AA; 55975 MW; E163773E0903D68C CRC64;

Query Match 10.8%; Score 86; DB 11; Length 488;

Best Local Similarity 33.6%; Pred. No. 0.14;

Matches 49; Conservative 10; Mismatches 39; Indels 48; Gaps 15;

QY 14 TELGGQYIKANSKFIGITELSSCH-GSEPCIIHR-GKPFGG----CHGSEPCIIHRG-KP 66

Db 256 TECG----KAFSRSSNLIEHQIHSGQPKYICKGKAFKGVSQLIHQ---LIHRGDKP 308

QY 67 FSSCHGSEPCIIHRGKPGGCHG-SEPCIIHRG-KPF--SSCHGSEPCIIHRGKPGG-- 120

Db 309 F-TCH-----EYKAFRGLSGLSQHORVHRGKPYECSEC-----GRAFGRA 350

QY 121 -----CHGSEPCIIH--RGKPFGR 137

Db 351 NLFKHQVVHGVVR-LQHRTRCKGFOR 375

Search completed: December 4, 2002, 12:52:58

Job time : 56.0875 secs

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 12:49:05 : Search time 67.9292 Seconds  
(without alignments)  
268.741 Million cell updates/sec

Title: US-09-362-731A-3

Perfect score: 793

Sequence: 1 DQYIKANSFIGITELGGQY.....FGGCHGSEPCIITHRGKPFSS 137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	793	100.0	137	21	AA1982634
2	160	20.2	32	21	AA1982636
3	111.5	14.1	136	22	AA198089
4	102.5	12.9	31	21	AA1982632
5	100	12.6	61	19	AA1982340
6	100	12.6	61	20	AA1980513
7	100	12.6	61	22	AA1981116
8	100	12.6	84	19	AA1982339
9	100	12.6	84	20	AA1980512
10	100	12.6	84	22	AA1981115

11	100	12.6	92	19	AA1982338
12	100	12.6	92	20	AA1980511
13	100	12.6	92	22	AA1981114
14	100	12.6	129	19	AA1982335
15	100	12.6	129	19	AA1982337
16	100	12.6	129	20	AA1980508
17	100	12.6	129	20	AA1980510
18	100	12.6	129	22	AA1981111
19	100	12.6	129	22	AA1981113
20	100	12.6	130	19	AA1982336
21	100	12.6	130	20	AA1980509
22	100	12.6	130	22	AA1981112
23	99	12.5	24	15	AA1981764
24	99	12.5	25	14	AA1983646
25	99	12.5	25	16	AA1982550
26	99	12.5	25	19	AA1981939
27	99	12.5	25	20	AA1980393
28	99	12.5	25	22	AA1981896
29	99	12.5	129	15	AA1984921
30	99	12.5	129	15	AA1984922
31	99	12.5	129	22	AA1980061
32	99	12.5	129	22	AA1987896
33	99	12.5	129	22	AA1980751
34	99	12.5	129	23	AB1986971
35	99	12.5	129	23	AB1986972
36	99	12.5	129	23	AB1986974
37	99	12.5	129	23	AB1986975
38	99	12.5	129	23	AB1986976
39	99	12.5	129	23	AB1986991
40	99	12.5	129	23	AB1986992
41	99	12.5	129	23	AB1986993
42	99	12.5	129	23	AB1986994
43	99	12.5	129	23	AB1986995
44	99	12.5	129	23	AB1986996
45	99	12.5	129	23	AB1986997

#### ALIGNMENTS

##### RESULT 1

AA1982634	AA1982634 standard; peptide: 137 AA.
ID	AA1982634 standard; peptide: 137 AA.
AC	AA1982634;
XX	07-AUG-2000 (first entry)
DT	Tetanus toxoid T cell epitopes and Der pII B cell epitopes peptide.
DE	T cell epitope; B cell epitope; allergen; antigenic;
KW	antiallergic; antiasthmatic; antiinflammatory; dermatological;
KW	immunosuppressive; vaccine; rhinitis; sinusitis; bronchial asthma;
KW	atopic dermatitis; acute urticaria; chronic urticaria;
KW	gastro-intestinal syndrome; food allergen; oro-pharyngeal syndrome;
KW	anaphylactic reaction; drug hypersensitivity; allergic reaction.
XX	Dermatophagoides pteronyssinus.
OS	Clostridium tetani.
OS	Synthetic.
XX	WO200006694-A2.
XX	10-FEB-2000.
XX	20-JUL-1999; 99WO-BE00092.
XX	30-JUL-1998; 98EP-0870167.
XX	(UNIO ) UCB SA.
XX	Saint-Remy J, Jacquemin M;



PN WO200072876-A2.  
 XX 07-DEC-2000.  
 XX 01-JUN-2000; 2000WO-US15239.  
 PF 01-JUN-1999; 99US-0137010.  
 XX (NEUR-) NEURALAB LTD.  
 XX Schenk DB;  
 XX WPI; 2001-070921/08.  
 XX Pharmaceutical composition comprising immunogen against amyloid  
 PT component such as fibril peptide or protein, or antibody against  
 PT amyloid component useful for treating amyloid diseases or amyloidoses -  
 XX Disclosure; Page 46; 140pp; English.  
 XX The invention relates to a novel pharmaceutical composition for  
 CC preventing or treating a disease characterised by amyloid fibril  
 CC deposits (amyloid plaques) in a patient. The pharmaceutical composition  
 CC comprises an agent that will induce an immune response against an amyloid  
 CC component, or an antibody or antibody fragment that binds to an amyloid  
 CC component. The invention also relates to a method for determining  
 CC the prognosis of a patient undergoing treatment for an amyloid disorder  
 CC which involves measuring a patient serum amount of immunoreactivity  
 CC against a selected amyloid component. A patient serum immunoreactivity  
 CC of at least four times a base line serum immunoreactivity control level  
 CC indicates a prognosis of improved status with respect to the disorder.  
 CC The pharmaceutical compositions of the invention are useful for treating  
 CC a wide variety of disorders characterised by amyloid fibril deposition in  
 CC a patient. Such disorders include Alzheimer's disease characterised by  
 CC amyloid beta peptide fibril deposits; type 2 diabetes characterised by  
 CC islet amyloid protein peptide (IAPP, amylin) fibrils; reactive systemic  
 CC amyloidosis associated with systemic inflammatory diseases (e.g.,  
 CC rheumatoid arthritis, osteomyelitis, tuberculosis) characterised by AA  
 CC fibrils derived from serum amyloid A protein (ApoSAA); systemic senile  
 CC amyloidosis and familial amyloid cardiomyopathy characterised by ATTR  
 CC fibrils derived from transthyretin (TTR); transmissible spongiform  
 CC encephalopathies (e.g. Creutzfeldt-Jakob disease, Kuru) characterised by  
 CC prion protein deposits; and beta-2-microglobulin deposits which form as  
 CC a result of long term haemodialysis treatment. The present sequence  
 CC represents an immunogenic fusion protein comprising an amyloid beta  
 CC peptide fused to a universal T-cell epitope which may be used in a  
 CC composition to treat or prevent Alzheimer's disease.  
 XX Sequence 136 AA;  
 SQ Query Match 14.1%; Score 111.5; DB 22; Length 136;  
 Best Local Similarity 38.0%; Pred. No. 0.00022;  
 Matches 30; Conservative 0; Mismatches 2; Indels 47; Gaps 1;  
 QY 2 QYIKANSKFIGITEL----- 16  
 DB 37 QYIKANSKFIGITELFNFTVSWFLRVPKVSASHLEDAEFRHDDAEFRHDDAEFRHDDAE 96  
 QY 17 --GGQYIKANSKFIGITEL 33  
 DB 97 FRHDDQYIKANSKFIGITEL 115  
 RESULT 4  
 ID AAY82632 standard; peptide; 31 AA.  
 XX AAY82632;  
 XX AAY82632;  
 DT 07-AUG-2000 (first entry)  
 XX Tetanus toxoid T cell epitope and Der pII B cell epitope peptide.  
 DE  
 XX

T cell epitope; B cell epitope; allergy; allergen; antigenic;  
 KW antiallergic; antiasthmatic; antiinflammatory; dermatological;  
 KW immunosuppressive; vaccine; rhinitis; sinusitis; bronchial asthma;  
 KW atopic dermatitis; acute urticaria; chronic urticaria;  
 KW gastro-intestinal syndrome; food allergen; oro-pharyngeal syndrome;  
 KW anaphylactic reaction; drug hypersensitivity; allergic reaction.  
 XX Dermatophagoides pteronyssinus.  
 OS Clostridium tetani.  
 OS Synthetic.  
 XX WO200006694-A2.  
 XX 10-FEB-2000.  
 PD 20-JUL-1999; 99WO-BE00092.  
 XX 30-JUL-1998; 98EP-0870167.  
 XX (UNIO ) UCB SA.  
 PA Saint-Remy J, Jacquemin M;  
 PI WPI; 2000-422470/36.  
 DR New compound for prevention and treatment of allergies comprises at  
 XX least one allergen antigenic determinant recognized by a B cell and at  
 XX least one antigenic determinant which does not trigger T cell  
 XX activation -  
 PS Claim 8; Page 35; 50pp; English.  
 XX The present invention describes a compound (I) for the prevention and/or  
 CC treatment of allergy. The compound comprises at least one allergen  
 CC antigenic determinant (i) recognised by a B cell or an antibody secreted  
 CC by a B cell of a non-atopic individual and at least one antigenic  
 CC determinant (ii) different from the allergen that triggers T cell  
 CC activation. (i) has antiallergic, antiasthmatic, antiinflammatory,  
 CC dermatological and immunosuppressive activities, and can be used in a  
 CC vaccine. (i) may be used in a pharmaceutical or cosmetic medicament to  
 CC treat and/or prevent allergies or a disease of allergic origin,  
 CC especially hypersensitivities. These include rhinitis, sinusitis,  
 CC bronchial asthma, atopic dermatitis, some forms of acute and chronic  
 CC urticaria, gastro-intestinal syndromes associated with the ingestion of  
 CC food allergens, oro-pharyngeal syndrome, anaphylactic reactions  
 CC associated with drug hypersensitivities and/or a mixture of these. The  
 CC use of (I) in the treatment of allergic conditions avoids the need for  
 CC drug treatment, which often causes undesirable side-effects. Also, prior  
 CC art drug therapies alleviate symptoms, but do not influence their  
 CC causes, however (I) actually combats the cause of an allergic reaction.  
 CC The present sequence represents a specifically claimed compound peptide  
 CC sequence from the present invention.  
 XX Sequence 31 AA;  
 SQ Query Match 12.9%; Score 102.5; DB 21; Length 31;  
 Best Local Similarity 60.5%; Pred. No. 0.00041;  
 Matches 23; Conservative 1; Mismatches 7; Indels 7; Gaps 1;  
 QY 2 QYIKANSKFIGITELGGQYIKANSKFIGITELSSCHGS 39  
 DB 1 QYIKANSKFIGITELGGHEIKR-----VLVPGCHGS 31  
 RESULT 5  
 ID AAW72340 standard; peptide; 61 AA.  
 XX AAW72340;  
 XX AAW72340;  
 DT 16-DEC-1998 (first entry)  
 XX Dermatophagoides Der f II clone MT16.  
 DE

```

XX genus Dermatophagoides; major protein allergen; T cell epitope;
KW Der p I; Der p II; Der f I; Der f II; house dust mite allergy.
XX
XX Dermatophagoides sp.
XX
XX US5820862-A.
XX
XX 13-OCT-1998.
XX
XX 07-JUN-1995; 95US-0482142.
XX
XX 19-MAY-1995; 95US-0445307.
XX
XX 14-APR-1994; 94US-0227772.
XX
XX 07-JUN-1995; 95US-0482142.
XX
XX (IMMU-) IMMULOGIC PHARM CORP.
XX
XX Chen X, Evans S, Franzen HM, Garman RD, Greenstein JL;
PI Kuo M, Rogers BL, Shaked Z;
XX
XX WPI; 1998-567590/48.
XX
XX Dermatophagoides allergen peptides - useful for treating house dust
PT mite allergy
XX
XX Disclosure; Fig 24; 155pp; English.
XX
XX The present invention describes peptides for treating sensitivity to
CC house dust mite allergens from the genus Dermatophagoides. Peptides
CC within the scope of the invention comprise at least one T cell epitope,
CC or preferably at least two T cell epitopes of a protein allergen
CC selected from the allergens Der p I, Der p II, Der f I, or Der f II.
CC The invention also describes modified peptides having similar or
CC enhanced therapeutic properties as the corresponding, naturally
CC occurring allergen, but having reduced side effects. The present
CC sequence represents a Der f II clone from the present invention.
XX
XX Sequence 61 AA;
SQ
Query Match 12.6%; Score 100; DB 19; Length 61;
Best Local Similarity 94.1%; Pred. No. 0.0016;
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 52 GCHGSEPCIIHRGKPFS 68
Db 20 GCHGSEPCIIHRGKPFT 36

RESULT 6
AAY50513
ID AAY50513 standard; Peptide; 61 AA.
XX
XX AAY50513;
XX
XX 25-JAN-2000 (first entry)
XX
XX Dermatophagoides sp major protein allergen fragment 13.
XX
XX Allergen; house dust mite; detection; sensitivity; T cell epitope;
KW screening; allergic disorder; asthma; rhinitis; ectopic dermatitis;
KW Der f I; Der p I; Der p II; Der f II.
XX
XX Dermatophagoides sp.
OS
XX
XX US5968526-A.
XX
XX 19-OCT-1999.
XX
XX 07-JUN-1995; 95US-0478572.
XX
XX 19-APR-1995; 95US-0445307.
XX
XX 14-APR-1994; 94US-0227772.
XX

```

```

PR 12-APR-1995; 95WO-US04481.
XX
XX (IMMU-) IMMULOGIC PHARM CORP.
XX
XX Garman RD, Greenstein JL, Rogers BL, Franzen HM, Shaked Z, Chen X;
PI Evans S, Kuo M;
XX
XX WPI; 1999-590385/50.
XX
XX Screening individuals for allergic reactions to T cell epitopes of
PT major allergens from house dust mites -
XX
XX Disclosure; Column 161-162; 158pp; English.
XX
XX This invention describes a novel method (I) for detecting whether an
CC individual is sensitive to Dermatophagoides (house dust mites). The
CC method involves detecting sensitivity to house dust mites in patients,
CC comprising combining a blood sample from the individual with 1 or more
CC isolated T cell epitopes of the protein allergens I and II ((DP I) and
CC (DP II)) from Dermatophagoides (house dust mites). 32 T cell epitopes
CC with varying, defined amino acids sequences (given in the specification)
CC may be used in (I). The sample and allergens are combined under
CC conditions appropriate for the binding of blood components with the
CC polypeptides. The extent of binding is then indicative of the
CC sensitivity of the patient to house dust mites. (I) may be used to screen
CC individuals for sensitivity to Dermatophagoides (house dust mites). The
CC house dust mite is a major cause of a variety of allergic disorders such
CC as asthma, rhinitis and ectopic dermatitis. AAY50360-Y50542 and
CC AAY50546-Y50555 represent house dust mite allergen peptide fragments
CC derived from Der p I, Der f II, Der f I and Der f II.
XX
XX Sequence 61 AA;
SQ
Query Match 12.6%; Score 100; DB 20; Length 61;
Best Local Similarity 94.1%; Pred. No. 0.0016;
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 52 GCHGSEPCIIHRGKPFS 68
Db 20 GCHGSEPCIIHRGKPFT 36

RESULT 7
AAU19116
ID AAU19116 standard; Protein; 61 AA.
XX
XX AAU19116;
XX
XX 04-DEC-2001 (first entry)
XX
XX House dust mite allergen Der f II variant MTL6.
XX
XX House dust mite; allergenic protein; Der p I; Der p II; Der f I;
KW Der f II; antiallergenic; immunostimulant; house dust mite allergy;
KW T-cell epitope.
XX
XX Dermatophagoides farinae.
OS
XX
XX US6268491-B1.
XX
XX 31-JUL-2001.
XX
XX 07-JUN-1995; 95US-0484296.
XX
XX 19-MAY-1995; 95US-0445307.
XX
XX 16-OCT-1991; 91US-0777859.
XX
XX 08-MAY-1992; 92US-0881396.
XX
XX 14-APR-1993; 93WO-US03471.
XX
XX 14-APR-1994; 94US-0227772.
XX
XX (IMMU-) IMMULOGIC PHARM CORP.
XX
XX Garman RD, Greenstein JL, Kuo M, Rogers BL, Franzen HM, Chen X;
PI

```



PI Evans S, Shaked Z;  
 XX WPI; 2001-549074/61.  
 XX  
 PT Peptides comprising T cell groups of the major allergens from  
 PT Dermatophagoides (house dust mites), useful for treating house dust  
 PT mite allergy in humans, and for diagnosing sensitivity to house dust  
 PT mite protein allergens -  
 XX  
 PS Disclosure; Figure 24; 158pp; English.  
 XX  
 CC The invention relates to an isolated peptide of the major protein  
 CC allergens of the genus Dermatophagoides, which comprises at least one T  
 CC cell group of a protein allergen from Der p (DP) I, DP II, Der f (DF) I  
 CC or DF II. The isolated peptide comprises at least two regions,  
 CC each region comprising at least one T cell group of a protein allergen  
 CC of the genus Dermatophagoides. The regions are derived from the same or  
 CC different protein allergens of the genus Dermatophagoides. The peptides  
 CC are useful for treating house dust mite allergy in humans. The peptides  
 CC are also useful for detecting or diagnosing sensitivity to house dust  
 CC mite protein allergens. The present peptides have similar or enhanced  
 CC therapeutic properties as the naturally-occurring allergen, but have  
 CC reduced side effects, and increased solubility and stability. The  
 CC present sequence represents an allergenic protein from Dermatophagoides  
 CC from which the T-cell epitope containing peptides are derived.  
 XX  
 CC  
 SO Sequence 61 AA;  
 Query Match 12.6%; Score 100; DB 22; Length 61;  
 Best Local Similarity 94.1%; Pred. No. 0.0016;  
 Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 52 GCHGSEPCIIHRGKPPS 68  
 Db 20 GCHGSEPCIIHRGKPPF 36  
 RESULT 8  
 AAW72339  
 ID AAW72339 standard; peptide; 84 AA.  
 XX  
 AC AAW72339;  
 XX  
 DT 16-DEC-1998 (first entry)  
 DE Dermatophagoides Der f II clone MT18.  
 XX  
 KW genus Dermatophagoides; major protein allergen; T cell epitope;  
 KW Der p I; Der p II; Der f I; Der f II; house dust mite allergy.  
 XX  
 OS Dermatophagoides sp.  
 XX  
 PN US5820862-A.  
 XX  
 PD 13-OCT-1998.  
 XX  
 PF 07-JUN-1995; 95US-0482142.  
 XX  
 PR 19-MAY-1995; 95US-0445307.  
 PR 14-APR-1994; 94US-0227772.  
 PR 07-JUN-1995; 95US-0482142.  
 XX  
 PA (IMMU-) IMMULOGIC PHARM CORP.  
 XX  
 PI Chen X, Evans S, Franzen HM, Garman RD, Greenstein JL;  
 PI Kuo M, Rogers BL, Shaked Z;  
 XX  
 XX WPI; 1998-567590/48.  
 XX Dermatophagoides allergen peptides - useful for treating house dust  
 PT mite allergy  
 XX  
 PS Disclosure; Fig 24; 155pp; English.

XX The present invention describes peptides for treating sensitivity to  
 CC house dust mite allergens from the genus Dermatophagoides. Peptides  
 CC within the scope of the invention comprise at least one T cell epitope,  
 CC or preferably at least two T cell epitopes of a protein allergen  
 CC selected from the allergens Der p I, Der p II, Der f I, or Der f II.  
 CC The invention also describes modified peptides having similar or  
 CC enhanced therapeutic properties as the corresponding, naturally  
 CC occurring allergen, but having reduced side effects. The present  
 CC sequence represents a Der f II clone from the present invention.  
 XX  
 CC  
 SQ Sequence 84 AA;  
 Query Match 12.6%; Score 100; DB 19; Length 84;  
 Best Local Similarity 94.1%; Pred. No. 0.0022;  
 Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 52 GCHGSEPCIIHRGKPPS 68  
 Db 20 GCHGSEPCIIHRGKPPF 36  
 RESULT 9  
 AAY50512  
 ID AAY50512 standard; Peptide; 84 AA.  
 XX  
 AC AAY50512;  
 XX  
 DT 25-JAN-2000 (first entry)  
 XX Dermatophagoides sp major protein allergen fragment 12.  
 DE Dermatophagoides sp major protein allergen fragment 12.  
 XX  
 KW Allergen; house dust mite; detection; sensitivity; T cell epitope;  
 KW screening; allergic disorder; asthma; rhinitis; ectopic dermatitis;  
 KW Der f I; Der p I; Der p II; Der f II.  
 XX  
 OS Dermatophagoides sp.  
 XX  
 PN US5968526-A.  
 XX  
 PD 19-OCT-1999.  
 XX  
 PF 07-JUN-1995; 95US-0478572.  
 XX  
 PR 19-MAY-1995; 95US-0445307.  
 PR 14-APR-1994; 94US-0227772.  
 PR 12-APR-1995; 95WO-US04481.  
 XX  
 PA (IMMU-) IMMULOGIC PHARM CORP.  
 XX  
 PI Garman RD, Greenstein JL, Rogers BL, Franzen HM, Shaked Z, Chen X;  
 PI Evans S, Kuo M;  
 XX  
 XX WPI; 1999-590385/50.  
 XX Screening individuals for allergic reactions to T cell epitopes of  
 PT major allergens from house dust mites -  
 PT  
 PS Disclosure; Column 161-162; 158pp; English.  
 XX  
 CC This invention describes a novel method (I) for detecting whether an  
 CC individual is sensitive to Dermatophagoides (house dust mites). The  
 CC method involves detecting sensitivity to house dust mites in patients,  
 CC comprising combining a blood sample from the individual with 1 or more  
 CC isolated T cell epitopes of the protein allergens I and II (DP I) and  
 CC (DP II) from Dermatophagoides (house dust mites). 32 T cell epitopes  
 CC with varying, defined amino acids sequences (given in the specification)  
 CC may be used in (I). The sample and allergens are combined under  
 CC conditions appropriate for the binding of blood components with the  
 CC polypeptides. The extent of binding is then indicative of the  
 CC sensitivity of the patient to house dust mites. (I) may be used to screen  
 CC individuals for sensitivity to Dermatophagoides (house dust mites). The  
 CC house dust mite is a major cause of a variety of allergic disorders such

CC as asthma, rhinitis and ectopic dermatitis. AAY50360-Y50542 and  
 CC AAY50546-Y50555 represent house dust mite allergen peptide fragments  
 CC derived from Der p I, Der f II, Der f I and Der f II.  
 XX  
 SQ Sequence 84 AA;

Query Match 12.6%; Score 100; DB 20; Length 84;  
 Best Local Similarity 94.1%; Pred. No. 0.0022;  
 Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 52 GCHGSEPCIIHRGKPPS 68  
 DB 20 GCHGSEPCIIHRGKPPS 36

RESULT 10  
 AAU19115  
 ID AAU19115 standard; Protein; 84 AA.

XX AAU19115;  
 XX  
 XX  
 XX  
 XX 04-DEC-2001 (first entry)  
 DE House dust mite allergen Der f II variant MT18.  
 XX  
 XX House dust mite; allergenic protein; Der p I; Der p II; Der f I;  
 KW Der f II; antiallergenic; immunostimulant; house dust mite allergy;  
 KW T-cell epitope.  
 XX

OS Dermatophagoides farinae.

XX US6268491-B1.  
 XX 31-JUL-2001.

XX 07-JUN-1995; 95US-0484296.

XX 19-MAY-1995; 95US-0445307.  
 PR 16-OCT-1991; 91US-0777859.  
 PR 08-MAY-1992; 92US-0881396.  
 PR 14-APR-1993; 93WO-US03471.  
 PR 14-APR-1994; 94US-0227772.

XX (IMMU-) IMMULOGIC PHARM CORP.

XX Garman RD, Greenstein JL, Kuo M, Rogers BL, Franzen HM, Chen X;  
 PI Evans S, Shaked Z;

XX WPI; 2001-549074/61.

XX Peptides comprising T cell groups of the major allergens from  
 PT Dermatophagoides (house dust mites), useful for treating house dust  
 PT mite allergy in humans, and for diagnosing sensitivity to house dust  
 PT mite protein allergens -

PS Disclosure; Figure 24; 158pp; English.

XX The invention relates to an isolated peptide of the major protein  
 CC allergens of the genus Dermatophagoides, which comprises at least one T  
 CC cell group of a protein allergen from Der p (DP) I, DP II, Der f (DF) I  
 CC or DF II. The isolated peptide comprises at least two regions,  
 CC each region comprising at least one T cell group of a protein allergen  
 CC of the genus Dermatophagoides. The regions are derived from the same or  
 CC different protein allergens of the genus Dermatophagoides. The peptides  
 CC are useful for treating house dust mite allergy in humans. The peptides  
 CC are also useful for detecting or diagnosing sensitivity to house dust  
 CC mite protein allergens. The present peptides have similar or enhanced  
 CC therapeutic properties as the naturally-occurring allergen, but have  
 CC reduced side effects, and increased solubility and stability. The  
 CC present sequence represents an allergenic protein from Dermatophagoides  
 CC from which the T-cell epitope containing peptides are derived.

XX Sequence 84 AA;

Query Match 12.6%; Score 100; DB 22; Length 84;  
 Best Local Similarity 94.1%; Pred. No. 0.0022;  
 Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 52 GCHGSEPCIIHRGKPPS 68  
 DB 20 GCHGSEPCIIHRGKPPS 36

RESULT 11  
 AAU72338  
 ID AAU72338 standard; peptide; 92 AA.

XX AAU72338;  
 XX  
 XX 16-DEC-1998 (first entry)  
 DT Dermatophagoides Der f II clone MT5.

XX genus Dermatophagoides; major protein allergen; T cell epitope;  
 KW Der p I; Der p II; Der f I; Der f II; house dust mite allergy.  
 XX

OS Dermatophagoides sp.

XX US5820862-A.

XX 13-OCT-1998.

XX 07-JUN-1995; 95US-0482142.

XX 19-MAY-1995; 95US-0445307.

PR 14-APR-1994; 94US-0227772.

PR 07-JUN-1995; 95US-0482142.

XX (IMMU-) IMMULOGIC PHARM CORP.

XX Chen X, Evans S, Franzen HM, Garman RD, Greenstein JL;  
 PI Kuo M, Rogers BL, Shaked Z;

XX WPI; 1998-567590/48.

XX Dermatophagoides allergen peptides - useful for treating house dust  
 PT mite allergy

XX Disclosure; Fig 24; 155pp; English.

XX The present invention describes peptides for treating sensitivity to  
 CC house dust mite allergens from the genus Dermatophagoides. Peptides  
 CC within the scope of the invention comprise at least one T cell epitope,  
 CC or preferably at least two T cell epitopes of a protein allergen  
 CC selected from the allergens Der p I, Der p II, Der f I, or Der f II.  
 CC The invention also describes modified peptides having similar or  
 CC enhanced therapeutic properties as the corresponding, naturally  
 CC occurring allergen, but having reduced side effects. The present  
 CC sequence represents a Der f II clone from the present invention.

XX Sequence 92 AA;

Query Match 12.6%; Score 100; DB 19; Length 92;  
 Best Local Similarity 94.1%; Pred. No. 0.0024;  
 Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 52 GCHGSEPCIIHRGKPPS 68  
 DB 20 GCHGSEPCIIHRGKPPS 36

RESULT 12  
 AAY50511  
 ID AAY50511 standard; Peptide; 92 AA.

XX AAY50511;

XX 25-JAN-2000 (first entry)  
 XX Dermatophagoides sp major protein allergen fragment 11.  
 DE Allergen: house dust mite; detection; sensitivity; T cell epitope;  
 XX screening; allergic disorder; asthma; rhinitis; ectopic dermatitis;  
 KW Der f I; Der p I; Der p II; Der f II.  
 KW Dermatophagoides sp.  
 OS US5968526-A.  
 PN 19-OCT-1999.  
 XX 07-JUN-1995; 95US-0478572.  
 XX 19-MAY-1995; 95US-0445307.  
 PR 14-APR-1994; 94US-0227772.  
 PR 12-APR-1995; 95WO-US04481.  
 XX (IMMU-) IMMULOGIC PHARM CORP.  
 PA Garman RD, Greenstein JL, Rogers BL, Franzen HM, Shaked Z, Chen X;  
 PI Evans S, Kuo M;  
 XX WPI; 1999-590385/50.  
 XX Screening individuals for allergic reactions to T cell epitopes of  
 PT major allergens from house dust mites -  
 PT Disclosure; Column 159-160; 158pp; English.  
 PS This invention describes a novel method (I) for detecting whether an  
 XX individual is sensitive to Dermatophagoides (house dust mites). The  
 CC method involves detecting sensitivity to house dust mites in patients,  
 CC comprising combining a blood sample from the individual with I or more  
 CC isolated T cell epitopes of the protein allergens I and II ((DP I) and  
 CC (DP II)) from Dermatophagoides (house dust mites). 32 T cell epitopes  
 CC with varying, defined amino acids sequences (given in the specification)  
 CC may be used in (I). The sample and allergens are combined under  
 CC conditions appropriate for the binding of blood components with the  
 CC polypeptides. The extent of binding is then indicative of the  
 CC sensitivity of the patient to house dust mites. (I) may be used to screen  
 CC individuals for sensitivity to Dermatophagoides (house dust mites). The  
 CC house dust mite is a major cause of a variety of allergic disorders such  
 CC as asthma, rhinitis and ectopic dermatitis. AAY50360-Y50542 and  
 CC AAY50546-Y50555 represent house dust mite allergen peptide fragments  
 CC derived from Der p I, Der f II, Der f I and Der f II.  
 XX Sequence 92 AA;  
 SQ Query Match 12.6%; Score 100; DB 20; Length 92;  
 Best Local Similarity 94.1%; Pred. No. 0.0024; Indels 0; Gaps 0;  
 Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 52 GCHGSEPCIHRGKPPF 68  
 Db 20 GCHGSEPCIHRGKPPF 36  
 RESULT 13  
 AAU19114  
 ID AAU19114 standard; Protein; 92 AA.  
 AC AAU19114;  
 XX 04-DEC-2001 (first entry)  
 DT House dust mite allergen Der f II variant MT5.  
 DE House dust mite; allergenic protein; Der p I; Der p II; Der f I;  
 XX Der f II; antiallergenic; immunostimulant; house dust mite allergy;  
 KW

KW T-cell epitope.  
 XX Dermatophagoides farinae.  
 OS US6268491-B1.  
 PN 31-JUL-2001.  
 XX 07-JUN-1995; 95US-0484296.  
 XX 19-MAY-1995; 95US-0445307.  
 PR 16-OCT-1991; 91US-0777859.  
 PR 08-MAY-1992; 92US-0881396.  
 PR 14-APR-1993; 93WO-US03471.  
 PR 14-APR-1994; 94US-0227772.  
 XX (IMMU-) IMMULOGIC PHARM CORP.  
 PA Garman RD, Greenstein JL, Kuo M, Rogers BL, Franzen HM, Chen X;  
 PI Evans S, Shaked Z;  
 XX WPI; 2001-549074/61.  
 XX Peptides comprising T cell groups of the major allergens from  
 PT Dermatophagoides (house dust mites), useful for treating house dust  
 PT mite allergy in humans, and for diagnosing sensitivity to house dust  
 PT mite protein allergens -  
 XX Disclosure; Figure 24; 158pp; English.  
 PS The invention relates to an isolated peptide of the major protein  
 XX allergens of the genus Dermatophagoides, which comprises at least one T  
 CC cell group of a protein allergen from Der p (DP) I, DP II, Der f (DF) I  
 CC or DF II. The isolated peptide comprises at least two regions,  
 CC each region comprising at least one T cell group of a protein allergen  
 CC of the genus Dermatophagoides. The regions are derived from the same or  
 CC different protein allergens of the genus Dermatophagoides. The peptides  
 CC are useful for treating house dust mite allergy in humans. The peptides  
 CC are also useful for detecting or diagnosing sensitivity to house dust  
 CC mite protein allergens. The present peptides have similar or enhanced  
 CC therapeutic properties as the naturally-occurring allergen, but have  
 CC reduced side effects, and increased solubility and stability. The  
 CC present sequence represents an allergenic protein from Dermatophagoides  
 CC from which the T-cell epitope containing peptides are derived.  
 XX Sequence 92 AA;  
 SQ Query Match 12.6%; Score 100; DB 22; Length 92;  
 Best Local Similarity 94.1%; Pred. No. 0.0024; Indels 0; Gaps 0;  
 Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 52 GCHGSEPCIHRGKPPF 68  
 Db 20 GCHGSEPCIHRGKPPF 36  
 RESULT 14  
 AAW72335  
 ID AAW72335 standard; peptide; 129 AA.  
 XX AC AAW72335;  
 XX 16-DEC-1998 (first entry)  
 DT Dermatophagoides Der f II clone pFL1.  
 DE genus Dermatophagoides; major protein allergen; T cell epitope;  
 KW Der p I; Der p II; Der f I; Der f II; house dust mite allergy.  
 XX Dermatophagoides sp.  
 OS US5820862-A.  
 PN

```

PD 13-OCT-1998.
XX
PF 07-JUN-1995; 95US-0482142.
XX
PR 19-MAY-1995; 95US-0445307.
PR 14-APR-1994; 94US-0227772.
XX 07-JUN-1995; 95US-0482142.
PA (IMMU-) IMMULOGIC PHARM CORP.
XX Chen X, Evans S, Franzen HM, Garman RD, Greenstein JL;
PI Kuo M, Rogers BL, Shaked Z;
XX WPI; 1998-567590/48.
DR
XX
PT Dermatophagoides allergen peptides - useful for treating house dust
PT mite allergy
XX
PS Disclosure; Fig 24; 155pp; English.
XX
CC The present invention describes peptides for treating sensitivity to
CC house dust mite allergens from the genus Dermatophagoides. Peptides
CC within the scope of the invention comprise at least one T cell epitope,
CC or preferably at least two T cell epitopes of a protein allergen
CC selected from the allergens Der p I, Der p II, Der f I, or Der f II.
CC The invention also describes modified peptides having similar or
CC enhanced therapeutic properties as the corresponding, naturally
CC occurring allergen, but having reduced side effects. The present
CC sequence represents a Der f II clone from the present invention.
XX
SQ Sequence 129 AA;

Query Match 12.6%; Score 100; DB 19; Length 129;
Best Local Similarity 94.1%; Pred. No. 0.0035;
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 52 GCHGSEPCIIHRGKPFS 68
DB 20 GCHGSEPCIIHRGKPFT 36

Search completed: December 4, 2002, 12:51:14
Job time : 68.1792 secs

RESULT 15
AAW72337
ID AAW72337 standard; peptide; 129 AA.
XX
AC AAW72337;
XX
DT 16-DEC-1998 (first entry)
XX
DE Dermatophagoides Der f II clone MT3.
XX
KW genus Dermatophagoides; major protein allergen; T cell epitope;
KW Der p I; Der p II; Der f I; Der f II; house dust mite allergy.
XX
OS Dermatophagoides sp.
XX
PN US5820862-A.
XX
PD 13-OCT-1998.
XX
PF 07-JUN-1995; 95US-0482142.
XX
PR 19-MAY-1995; 95US-0445307.
PR 14-APR-1994; 94US-0227772.
XX 07-JUN-1995; 95US-0482142.
XX
PA (IMMU-) IMMULOGIC PHARM CORP.
XX
PI Chen X, Evans S, Franzen HM, Garman RD, Greenstein JL;
PI Kuo M, Rogers BL, Shaked Z;
XX WPI; 1998-567590/48.
XX

```

```

PT Dermatophagoides allergen peptides - useful for treating house dust
PT mite allergy
XX
PS Disclosure; Fig 24; 155pp; English.
XX
CC The present invention describes peptides for treating sensitivity to
CC house dust mite allergens from the genus Dermatophagoides. Peptides
CC within the scope of the invention comprise at least one T cell epitope,
CC or preferably at least two T cell epitopes of a protein allergen
CC selected from the allergens Der p I, Der p II, Der f I, or Der f II.
CC The invention also describes modified peptides having similar or
CC enhanced therapeutic properties as the corresponding, naturally
CC occurring allergen, but having reduced side effects. The present
CC sequence represents a Der f II clone from the present invention.
XX
SQ Sequence 129 AA;

Query Match 12.6%; Score 100; DB 19; Length 129;
Best Local Similarity 94.1%; Pred. No. 0.0035;
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 52 GCHGSEPCIIHRGKPFS 68
DB 20 GCHGSEPCIIHRGKPFT 36

Search completed: December 4, 2002, 12:51:14
Job time : 68.1792 secs

```

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model  
Run on: December 4, 2002, 12:53:35 : Search time 12.5583 Seconds  
(without alignments)  
177.189 Million cell updates/sec

Title: US-09-362-731A-3  
Perfect score: 793  
Sequence: 1 DQYIKANSKFITIGLGGQY.....FGGCHGSEPCIHRCKPFSR 137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	99	12.5	145	10	US-09-949-889-3
2	93	11.7	129	10	US-09-949-889-4
3	82	10.3	31	10	US-09-983-019-6
4	82	10.3	46	10	US-09-848-834A-19
5	80	10.1	28	10	US-09-848-834A-11
6	78	9.8	383	10	US-09-205-658-105
7	78	9.8	383	10	US-09-844-353A-105
8	77.5	9.8	1801	10	US-09-938-275-8
9	76	9.6	31	10	US-09-943-548-2
10	75	9.5	31	10	US-09-983-019-5
11	75	9.5	1701	9	US-09-963-959-2
12	74.5	9.4	311	10	US-09-864-761-34333
13	74.5	9.4	340	10	US-09-864-761-35894
14	74.5	9.4	1798	10	US-09-938-275-9
15	74	9.3	15	10	US-09-862-849-2
16	74	9.3	16	10	US-09-848-834A-2
17	74	9.3	29	10	US-09-732-754-1
18	74	9.3	29	10	US-09-983-019-8
19	74	9.3	31	10	US-09-848-834A-15

20	74	9.3	50	10	US-09-943-548-8	Sequence 8, Appli
21	73.5	9.3	557	10	US-09-864-761-34705	Sequence 34705, A
22	73	9.2	179	9	US-09-981-876-177	Sequence 177, App
23	73	9.2	310	10	US-09-764-853-542	Sequence 542, App
24	73	9.2	310	10	US-09-764-858-180	Sequence 180, App
25	73	9.2	413	10	US-09-925-299-977	Sequence 977, App
26	73	9.2	1055	10	US-09-855-722-2	Sequence 2, Appli
27	73	9.2	1212	10	US-09-855-722-3	Sequence 3, Appli
28	73	9.2	1238	10	US-09-855-722-5	Sequence 5, Appli
29	73	9.2	1238	10	US-09-944-849-4	Sequence 4, Appli
30	73	9.2	1523	12	US-10-011-064-2	Sequence 2, Appli
31	73	9.2	1523	12	US-10-052-586-290	Sequence 290, App
32	72	9.1	207	9	US-09-976-736-13	Sequence 13, Appli
33	72	9.1	233	10	US-09-216-393-110	Sequence 110, App
34	72	9.1	757	10	US-09-919-603-3	Sequence 3, Appli
35	72	9.1	5405	10	US-09-922-217-1116	Sequence 1116, Ap
36	71	9.0	31	10	US-09-983-019-3	Sequence 3, Appli
37	71	9.0	249	10	US-09-764-864-1445	Sequence 1445, Ap
38	71	9.0	295	10	US-09-764-864-1016	Sequence 1016, Ap
39	71	9.0	845	10	US-09-898-570-12	Sequence 12, Appli
40	71	9.0	974	10	US-09-898-570-14	Sequence 14, Appli
41	70.5	8.9	997	10	US-09-747-371-3	Sequence 3, Appli
42	70.5	8.9	999	10	US-09-747-371-2	Sequence 2, Appli
43	70	8.8	29	10	US-09-983-019-9	Sequence 9, Appli
44	70	8.8	50	10	US-09-943-548-9	Sequence 9, Appli
45	70	8.8	342	10	US-09-864-761-36708	Sequence 36708, A

ALIGNMENTS

RESULT 1  
US-09-949-889-3  
; Sequence 3, Application US/09949889  
; Patent No. US20020054881A1  
; GENERAL INFORMATION:  
; APPLICANT: CONSIGLIO NAZIONALE DELLE RICERCHE  
; TITLE OF INVENTION: VARIANTS OF ALLERGENIC PROTEINS OF THE GROUP 2 OF  
; TITLE OF INVENTION: DERMATOPHAGOIDES  
; FILE REFERENCE: Cons Naz Ric  
; CURRENT APPLICATION NUMBER: US/09/949,889  
; CURRENT FILING DATE: 2001-09-12  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 145  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus  
US-09-949-889-3

Query Match 12.5%; Score 99; DB 10; Length 145;  
Best Local Similarity 100.0%; Pred. No. 0.00066;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	52	GCHGSEPCIHRGKPF	67
Db	36	GCHGSEPCIHRGKPF	51

RESULT 2  
US-09-949-889-4  
; Sequence 4, Application US/09949889  
; Patent No. US20020054881A1  
; GENERAL INFORMATION:  
; APPLICANT: CONSIGLIO NAZIONALE DELLE RICERCHE  
; TITLE OF INVENTION: VARIANTS OF ALLERGENIC PROTEINS OF THE GROUP 2 OF  
; TITLE OF INVENTION: DERMATOPHAGOIDES  
; FILE REFERENCE: Cons Naz Ric  
; CURRENT APPLICATION NUMBER: US/09/949,889  
; CURRENT FILING DATE: 2001-09-12  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4



```

Query Match      10.1%; Score 80; DB 10; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 16; Conservative 0; Mismatches 0; Indels

Qy  2 QYKANSKFIGITELG 17
Db  1 QYKANSKFIGITELG 16

```

```

RESULT 6
us-09-205-658-105
; Sequence 105, Application US/09205658
; Patent No. US20010029617A1
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary
; APPLICANT: Ogg, Scott
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; TITLE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
; FILE REFERENCE: 00786/351004
; CURRENT APPLICATION NUMBER: US/09/205, 658
; CURRENT FILING DATE: 1998-12-03
; EARLIER APPLICATION NUMBER: 08/857, 076
; EARLIER FILING DATE: 1997-05-15
; EARLIER APPLICATION NUMBER: 08/888, 534
; EARLIER FILING DATE: 1997-07-07
; EARLIER APPLICATION NUMBER: US98/10080
; EARLIER FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 328
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 105
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-205-658-105

```

```

Query Match          9.8%; Score 78; DB 10; Length 383
Best Local Similarity 23.3%; Pred. No. 0.31;
Matches 35; Conservative 13; Mismatches 38; Indels
-----
QY 30 ITELSSCHGSEPCIHRGKPPGCG---HGSEPCIIHRGKPF-----
Db          : : : : : : : : : : : : : : : : : :
QY 99 IDEHTCC---SQDCL-----GGCVIDKNESCISCNVNFSNMCMDSCPFG
Db          : : : : : : : : : : : : : : : : : :
QY 70 CHGSEPCI-----IHRGKPPGCGHSGSEPCIHRGKPFSSCHGS---EPC
Db          : : : : : : : : : : : : : : : : : :
QY 150 CVTANECITLTKFTNSVSYGPIYNG-----QCITHCTPGYQKSENKRMCEPC
Db          : : : : : : : : : : : : : : : : : :
QY 111 -----IIHRGKPPGCG---HGSEPCII 129
Db          : : : : : : : : : : : : : : : : : :
QY 205 ECSSGLIDSLERAREFHGCTIITGTPEPLTI 234
Db          : : : : : : : : : : : : : : : : : :

```

```

RESULT 7
US-09-844-353A-105
; Sequence 105, Application US/09844353A
; Patent No. US20020037585A1
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary
; APPLICANT: Kimura, Koutarou
; APPLICANT: Patterson, Garth
; APPLICANT: Ogg, Scott
; APPLICANT: Paradis, Suzanne
; APPLICANT: Tissenbaum, Heidi
; APPLICANT: Morris, Jason
; APPLICANT: Koweek, Allison
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; TITILE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
; FILE REFERENCE: 00786/351005
; CURRENT APPLICATION NUMBER: US/09/844,353A
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 08/857,076
; PRIOR FILING DATE: 1997-05-15
; NUMBER OF SEQ ID NOS: 114

```

```

Query Match          10.1%; Score 80; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 16; Conservative 0; Mismatches 0; Indels

Qy      2 QYIKANSKFIFGITELG 17
        |||||
Db      1 QYIKANSKFIFGITELG 16

RESULT 6
US-09-205-658-105
; Sequence 105, Application US/09205658
; Patent No. US20010029617A1
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary
; APPLICANT: Ogg, Scott
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; TITLE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
; FILE REFERENCE: 00786/351004
; CURRENT APPLICATION NUMBER: US/09/205,658
; CURRENT FILING DATE: 1998-12-03
; EARLIER APPLICATION NUMBER: 08/857,076

```

```

; EARLIER FILING DATE: 1997-07-07
; EARLIER APPLICATION NUMBER: US98/10080
; EARLIER FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 328
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 105
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-205-658-105

Query Match          9.8%;   Score 78;   DB 10;   Length 383
Best Local Similarity 23.3%;   Pred. No. 0.31;
Matches 35;   Conservative 13;   Mismatches 38;   Indels

QY 30 ITELSSCHGSEPCIHRGKPFPGC-----HGSEPCIHRGKGF-----
      | | | | | | | | | | | | | | | | | | | | | | | |
Db 99 IDEHTCC--SQDCL-----GSCVDKNGNESCISRNVSFNICMDSCPKG

QY 70 CHGSEPCI-----IHRKPFPGCHGSEPCIHRGKPFSSCHGS-----BPC
      | | | | | | | | | | | | | | | | | | | | | | | |
Db 150 CVTANECHITLTKFETNSVSGIPYNG-----QCITHCPTGQKSENKRMCEPC
      | | | | | | | | | | | | | | | | | | | | | | | |

```

```

QY  111  -----IIHRCKPFGC---HGSEPCI 129
      : | : | | |
      : | : | | |
Db  205  ECSSGLIDSLERAREFHGCTIITGTEPLT 234

```

RESULT 7  
 US-09-844-353A-105  
 ; Sequence.105, Application US/09844353A  
 ; Patent NO. US20020037585A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ruvkun, Gary  
 ; APPLICANT: Kimura, Koutarou  
 ; APPLICANT: Patterson, Garth  
 ; APPLICANT: Ogg, Scott  
 ; APPLICANT: Paradis, Suzanne

```

: APPLICANT: Kowcek, Allison
:
: TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
:
: TITLE OF INVENTION: IMPAIRED GLOUCESTER TOLERANCE CONDITIONS
:
: FILE REFERENCE: 00786/351005
:
: CURRENT APPLICATION NUMBER: US/09/844,353A
:
: CURRENT FILING DATE: 2001-04-27
:
: PRIOR APPLICATION NUMBER: US 08/857,076
:
: PRIOR FILING DATE: 1997-05-15
:
: NUMBER OF SEQ ID NOS: 114
:

```





```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35894
; LENGTH: 340
; TYPE: prt
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AF166490.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN BT47A, SIGNAL = 9.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 6.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3
; OTHER INFORMATION: EST_HUMAN HIT: AFI50239.1, EVALUAE 1.00e-91
; OTHER INFORMATION: SWISSPROT HIT: P52742, EVALUAE 1.00e-118
US-09-864-761-35894

Query Match          9.4%; Score 74.5; DB 10; Length 340;
Best Local Similarity 25.6%; Pred.No. 0.64;
Matches 42; Conservative 9; Mismatches 38; Indels 75; Gaps ?

Qy   43 IHRGKPFGGCH-----GSEPCIIHR-GKPFS--CHG 72
      :| | : |||           |:| |:| |||| 
Db    51 MIHTGERPHGCGLCGKAFTHCSDLRKHERTHTGKPYGCLCKAFKSNNLRHHEIMT 110
       :| | |             |:| |:| |||| 

Qy   73 SEPC-IIHR-GKPPGCC-----GSEP-CIIRHGKPFSSC-----HGSE 108
      | | | | |         -:| -:| |:: |:| |:: |:| 
Db    11 REKAQCILCHLCGAFTHCSDLRKHERTHLGDGPYGCLL-CGAKFSCSYLRQHETHNGEK 169
        | | | | |         -|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|
Qy   109 PCIHR-RGKPGGC-----HGSEPCIHRGKFPES 136
       | | | | |         |-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|
Db    170 PYEHLGCAFHCSHLRQHSHNGKEPKGHCLC----GKAT 209
        | | | | |         |-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|

RESULT 14
US-09-938-275-9
; Sequence 9, Application US/09938275
; Patent No. US20020111309A1
; GENERAL INFORMATION:
; APPLICANT: Gerardo Castillo
; APPLICANT: Alan Snow
; TITLE OF INVENTION: Therapeutic and Diagnostic Applications
; FILE REFERENCE: Of Laminin and Laminin-Derived Protein Fragments
; CURRENT APPLICATION NUMBER: US/09/938,275
; CURRENT FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 11
```

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1798
; TYPE: PRT
; ORGANISM: Homo Sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Swissprot P55268
; DATABASE ENTRY DATE: 1996-10-01
US-09-938-275-9

Query Match          9.4%; Score 74.5; DB 10; Length 1798;
Best Local Similarity 25.9%; Pred. No. 4;
Matches 36; Conservative 7; Mismatches 45; Indels 51; Gaps 10;

QY 27 FTGITELSSCH-----GSEP-----CIIHRGKPFGGCHGSEPCIIHRGKP----- 66
   || || || || || || || || || || || || || || || || || || || || ||
Db 1026 FHQAARQCPCHCTNLTNPQCPSPDQC--HCDPSSGQC---PCLPNVQGPSCDRC 1079
   || || || || || || || || || || || || || || || || || || || || ||

QY 67 -----FSSCHGSEPCIIH-----RGKP-----FGGCH-----GSEPCI-----IHRGKPFSS 103
   : || || || || || || || || || || || || || || || || || || || || ||
Db 1080 APFNWNLTSHGCGCPACHPSPARGPTCNFTGQCHCRAGFGGRTCEQCELHWGDPGLQ 1139
   || || || || || || || || || || || || || || || || || || || || ||

QY 104 CHGSEPCIIHRGKPGGGCH 122
   || || || || || || || || || || || || || || || || || || || || ||
Db 1140 CHACD--CDSRGIDTPQCH 1156
   || || || || || || || || || || || || || || || || || || || || ||

RESULT 15
US-09-862-849-2
; Sequence 2, Application US/09862849
; Patent No. US20020013274A1
; GENERAL INFORMATION:
; APPLICANT: Sudhir Paul
; APPLICANT: Larry J. Smith
; APPLICANT: Gennady Gololobov
; TITLE OF INVENTION: Methods for Identifying Inducers and Inhibitors of Proteolytic
; FILE OF INVENTION: Antibodies, Compositions and Their Uses
; FILE REFERENCE: UNMC 63123 DIV
; CURRENT APPLICATION NUMBER: US/09/862,849
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: US 09/046,373
; PRIOR FILING DATE: 1998-03-23
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-09-862-849-2

Query Match          9.3%; Score 74; DB 10; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QYIKANSKFIGITEL 16
   || || || || || || || || || || || || || || || || || || || || ||
Db 1 QYIKANSKFIGITEL 15
   || || || || || || || || || || || || || || || || || || || || ||

Search completed: December 4, 2002, 13:05:29
Job time : 13.5583 secs
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 12:49:07 ; Search time 22.8333 Seconds  
(without alignments)  
176.538 Million cell updates/sec

Title: US-09-362-731A-3

Perfect score: 793

Sequence: 1 DQYIRANSFKFIGHTELGGY.....FGCHGSEPCIIHRKPFESR 137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents\_AA:\*

1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*

2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*

3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*

4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*

5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*

6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	12.6	61	2	US-08-482-142-162
2	100	12.6	61	2	US-08-478-572-162
3	100	12.6	61	4	US-08-484-296-162
4	100	12.6	84	2	US-08-482-142-161
5	100	12.6	84	2	US-08-478-572-161
6	100	12.6	84	4	US-08-484-296-161
7	100	12.6	92	2	US-08-482-142-160
8	100	12.6	92	2	US-08-478-572-160
9	100	12.6	92	4	US-08-484-296-160
10	100	12.6	129	2	US-08-482-142-157
11	100	12.6	129	2	US-08-482-142-159
12	100	12.6	129	2	US-08-478-572-157
13	100	12.6	129	2	US-08-478-572-159
14	100	12.6	129	4	US-08-484-296-157
15	100	12.6	129	4	US-08-484-296-159
16	100	12.6	130	2	US-08-482-142-158
17	100	12.6	130	2	US-08-478-572-158
18	100	12.6	130	4	US-08-484-296-158
19	99	12.5	25	2	US-08-482-142-42
20	99	12.5	25	2	US-08-478-572-42
21	99	12.5	25	4	US-08-484-296-42
22	99	12.5	25	5	PCT-US95-04481-19
23	99	12.5	35	3	US-08-460-040-7
24	99	12.5	129	1	US-07-945-288-12
25	99	12.5	129	1	US-08-462-831-12
26	99	12.5	129	1	US-08-461-809-12
27	99	12.5	129	1	US-08-461-441-12

28 99 12.5 129 5 PCT-US93-08518-12 Sequence 12, Appli  
29 99 12.5 145 3 US-08-460-040-6 Sequence 6, Appli  
30 99 12.5 146 1 US-07-945-288-4 Sequence 4, Appli  
31 99 12.5 146 1 US-08-462-831-4 Sequence 4, Appli  
32 99 12.5 146 1 US-08-461-809-4 Sequence 4, Appli  
33 99 12.5 146 1 US-08-461-441-4 Sequence 4, Appli  
34 99 12.5 146 2 US-08-482-142-4 Sequence 4, Appli  
35 99 12.5 146 2 US-08-478-572-4 Sequence 4, Appli  
36 99 12.5 146 4 US-08-484-296-4 Sequence 4, Appli  
37 99 12.5 146 5 PCT-US93-08518-4 Sequence 4, Appli  
38 97 12.2 129 1 US-07-945-288-8 Sequence 8, Appli  
39 97 12.2 129 1 US-07-945-288-13 Sequence 13, Appli  
40 97 12.2 129 1 US-08-462-831-8 Sequence 8, Appli  
41 97 12.2 129 1 US-08-462-831-13 Sequence 13, Appli  
42 97 12.2 129 1 US-08-461-809-8 Sequence 8, Appli  
43 97 12.2 129 1 US-08-461-809-13 Sequence 13, Appli  
44 97 12.2 129 1 US-08-461-441-8 Sequence 8, Appli  
45 .97 12.2 129 1 US-08-461-441-13 Sequence 13, Appli

#### ALIGNMENTS

RESULT 1  
US-08-482-142-162  
; Sequence 162, Application US/08482142  
; Patent No. 5820862  
; GENERAL INFORMATION:  
; APPLICANT: Garman, Richard  
; APPLICANT: Greenstein, Julia  
; APPLICANT: Kuo, Mei-chang  
; APPLICANT: Rogers, Bruce  
; APPLICANT: Franzen, Henry  
; APPLICANT: Chen, Xian  
; APPLICANT: Evans, Sean  
; APPLICANT: Shaked, Ze'ev  
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
; NUMBER OF SEQUENCES: 207  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
; STREET: 610 LINCOLN STREET  
; CITY: WALTHAM  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/482,142  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/445,307  
; FILING DATE: 07 June 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CRAIG, ANNE I.  
; REGISTRATION NUMBER: 32,976  
; REFERENCE/DOCKET NUMBER: 017.605  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 466-6000  
; TELEFAX: (617) 466-6040  
; INFORMATION FOR SEQ ID NO: 162:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 61 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: internal  
US-08-482-142-162



```
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,142
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,307
; FILING DATE: 07 June 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAIG, ANNE I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: 017.6US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6040
; INFORMATION FOR SEQ ID NO: 161:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
;
; US-08-482-142-161
;
; Query Match 12.6%; Score 100; DB 2; Length 84;
; Best Local Similarity 94.1%; Pred. No. 0.00016;
; Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
;
; QY 52 GCHGSEPCIIHRGKPF 68
;      |||||
; DB 20 GCHGSEPCIIHRGKPF 36
;
; RESULT 5
; US-08-478-572-161
; Sequence 161, Application US/08478572
; Patent No. 5968526
;
; GENERAL INFORMATION:
; APPLICANT: Garman, Richard
; APPLICANT: Greenstein, Julia
; APPLICANT: Kuo, Mei-chang
; APPLICANT: Rogers, Bruce
; APPLICANT: Franzen, Henry
; APPLICANT: Chen, Xian
; APPLICANT: Evans, Sean
; APPLICANT: Shaked, Ze'ev
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS
; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,142
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,307
; FILING DATE: 07 June 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAIG, ANNE I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: 017.6US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6040
; INFORMATION FOR SEQ ID NO: 161:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
;
; US-08-482-142-161
```

```
;
; APPLICATION NUMBER: US/08/478,572
; FILING DATE: 07-June-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/445,307
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAIG, ANNE I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: 017.6US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6040
; INFORMATION FOR SEQ ID NO: 161:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
;
; US-08-478-572-161
;
; Query Match 12.6%; Score 100; DB 2; Length 84;
; Best Local Similarity 94.1%; Pred. No. 0.00016;
; Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
;
; QY 52 GCHGSEPCIIHRGKPF 68
;      |||||
; DB 20 GCHGSEPCIIHRGKPF 36
;
; RESULT 6
; US-08-484-296-161
; Sequence 161, Application US/08484296
; Patent No. 6268491
;
; GENERAL INFORMATION:
; APPLICANT: Garman, Richard
; APPLICANT: Greenstein, Julia
; APPLICANT: Kuo, Mei-chang
; APPLICANT: Rogers, Bruce
; APPLICANT: Franzen, Henry
; APPLICANT: Chen, Xian
; APPLICANT: Evans, Sean
; APPLICANT: Shaked, Ze'ev
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS
; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,296
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/445,307
; FILING DATE: 07 June 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAIG, ANNE I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: 017.6US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6040
```

INFORMATION FOR SEQ ID NO: 161:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 84 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-484-296-161

Query Match 12.6%; Score 100; DB 4; Length 84;  
Best Local Similarity 94.1%; Pred. No. 0.00016;  
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 52 GCHGSEPCIHRGKPF 68  
Db 20 GCHGSEPCIHRGKPF 36

RESULT 7  
US-08-482-142-160  
; Sequence 160, Application US/08482142  
; Patent No. 5820862  
; GENERAL INFORMATION:  
; APPLICANT: Garman, Richard  
; APPLICANT: Greenstein, Julia  
; APPLICANT: Kuo, Mei-chang  
; APPLICANT: Rogers, Bruce  
; APPLICANT: Franzen, Henry  
; APPLICANT: Chen, Xian  
; APPLICANT: Evans, Sean  
; APPLICANT: Shaked, Ze'ev  
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
; NUMBER OF SEQUENCES: 207  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
; STREET: 610 LINCOLN STREET  
; CITY: WALTHAM  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02154

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,142  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/445,307  
FILING DATE: 07 June 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: CRAIG, ANNE I.  
REGISTRATION NUMBER: 32,976  
REFERENCE/DOCKET NUMBER: 017.6US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 466-6000  
TELEFAX: (617) 466-6040  
INFORMATION FOR SEQ ID NO: 160:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 92 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-482-142-160

Query Match 12.6%; Score 100; DB 2; Length 92;  
Best Local Similarity 94.1%; Pred. No. 0.00017;  
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 52 GCHGSEPCIHRGKPF 68  
Db 20 GCHGSEPCIHRGKPF 36

RESULT 8  
US-08-478-572-160  
; Sequence 160, Application US/08478572  
; Patent No. 5968526  
; GENERAL INFORMATION:  
; APPLICANT: Garman, Richard  
; APPLICANT: Greenstein, Julia  
; APPLICANT: Kuo, Mei-chang  
; APPLICANT: Rogers, Bruce  
; APPLICANT: Franzen, Henry  
; APPLICANT: Chen, Xian  
; APPLICANT: Evans, Sean  
; APPLICANT: Shaked, Ze'ev  
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
; NUMBER OF SEQUENCES: 207  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
; STREET: 610 LINCOLN STREET  
; CITY: WALTHAM  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02154

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/478,572  
FILING DATE: 07-June-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/445,307  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: CRAIG, ANNE I.  
REGISTRATION NUMBER: 32,976  
REFERENCE/DOCKET NUMBER: 017.6US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 466-6000  
TELEFAX: (617) 466-6040  
INFORMATION FOR SEQ ID NO: 160:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 92 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-478-572-160

Query Match 12.6%; Score 100; DB 2; Length 92;  
Best Local Similarity 94.1%; Pred. No. 0.00017;  
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 52 GCHGSEPCIHRGKPF 68  
Db 20 GCHGSEPCIHRGKPF 36

RESULT 9  
US-08-484-296-160  
; Sequence 160, Application US/08484296  
; Patent No. 6268491  
; GENERAL INFORMATION:  
; APPLICANT: Garman, Richard  
; APPLICANT: Greenstein, Julia  
; APPLICANT: Kuo, Mei-chang

APPLICANT: Rogers, Bruce  
APPLICANT: Franzen, Henry  
APPLICANT: Chen, Xian  
APPLICANT: Evans, Sean  
APPLICANT: Shaked, Ze'ev  
TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
NUMBER OF SEQUENCES: 207  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
STREET: 610 LINCOLN STREET  
CITY: WALTHAM  
STATE: MA  
COUNTRY: USA  
ZIP: 02154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,296  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/445,307  
FILING DATE: 07 June 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: CRAIG, ANNE I.  
REGISTRATION NUMBER: 32,976  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 466-6000  
TELEFAX: (617) 466-6040  
INFORMATION FOR SEQ ID NO: 160:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 92 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-484-296-160  
  
Query Match 12.6%; Score 100; DB 4; Length 92;  
Best Local Similarity 94.1%; Pred. No. 0.00017;  
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 52 GCHGSEPCIIHRGKPF 68  
Db 20 GCHGSEPCIIHRGKPF 36  
|||||  
  
RESULT 10  
US-08-482-142-157  
Sequence 157, Application US/08482142  
Patent No. 5820862  
GENERAL INFORMATION:  
APPLICANT: Garman, Richard  
APPLICANT: Greenstein, Julia  
APPLICANT: Kuo, Mei-chang  
APPLICANT: Rogers, Bruce  
APPLICANT: Franzen, Henry  
APPLICANT: Chen, Xian  
APPLICANT: Evans, Sean  
APPLICANT: Shaked, Ze'ev  
TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
NUMBER OF SEQUENCES: 207  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
STREET: 610 LINCOLN STREET  
CITY: WALTHAM  
STATE: MA

COUNTRY: USA  
ZIP: 02154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,142  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/445,307  
FILING DATE: 07 June 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: CRAIG, ANNE I.  
REGISTRATION NUMBER: 32,976  
REFERENCE/DOCKET NUMBER: 017.60S  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 466-6000  
TELEFAX: (617) 466-6040  
INFORMATION FOR SEQ ID NO: 157:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 129 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-482-142-157  
  
Query Match 12.6%; Score 100; DB 2; Length 129;  
Best Local Similarity 94.1%; Pred. No. 0.00026;  
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 52 GCHGSEPCIIHRGKPF 68  
Db 20 GCHGSEPCIIHRGKPF 36  
|||||  
  
RESULT 11  
US-08-482-142-159  
Sequence 159, Application US/08482142  
Patent No. 5820862  
GENERAL INFORMATION:  
APPLICANT: Garman, Richard  
APPLICANT: Greenstein, Julia  
APPLICANT: Kuo, Mei-chang  
APPLICANT: Rogers, Bruce  
APPLICANT: Franzen, Henry  
APPLICANT: Chen, Xian  
APPLICANT: Evans, Sean  
APPLICANT: Shaked, Ze'ev  
TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
NUMBER OF SEQUENCES: 207  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
STREET: 610 LINCOLN STREET  
CITY: WALTHAM  
STATE: MA  
COUNTRY: USA  
ZIP: 02154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,142  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/445,307

```
; FILING DATE: 07 June 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAIG, ANNE I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: 017.60S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6040
; INFORMATION FOR SEQ ID NO: 159:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 129 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-08-482-142-159

Query Match 12.6%; Score 100; DB 2; Length 129;
Best Local Similarity 94.1%; Pred. No. 0.00026;
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 52 GCHGSEPCIIHRGKPF 68
Db 20 GCHGSEPCIIHRGKPF 36

RESULT 12
US-08-478-572-157
; Sequence 157, Application US/08478572
; Patent No. 5968526
; GENERAL INFORMATION:
; APPLICANT: Garman, Richard
; APPLICANT: Greenstein, Julia
; APPLICANT: Kuo, Mei-chang
; APPLICANT: Rogers, Bruce
; APPLICANT: Franzen, Henry
; APPLICANT: Chen, Xian
; APPLICANT: Evans, Sean
; APPLICANT: Shaked, Ze'ev
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS
; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,572
; FILING DATE: 07-June-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/445,307
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAIG, ANNE I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: 017.60S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6040
; INFORMATION FOR SEQ ID NO: 157:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 129 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-08-478-572-157

Query Match 12.6%; Score 100; DB 2; Length 129;
Best Local Similarity 94.1%; Pred. No. 0.00026;
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 52 GCHGSEPCIIHRGKPF 68
Db 20 GCHGSEPCIIHRGKPF 36

RESULT 13
US-08-478-572-159
; Sequence 159, Application US/08478572
; Patent No. 5968526
; GENERAL INFORMATION:
; APPLICANT: Garman, Richard
; APPLICANT: Greenstein, Julia
; APPLICANT: Kuo, Mei-chang
; APPLICANT: Rogers, Bruce
; APPLICANT: Franzen, Henry
; APPLICANT: Chen, Xian
; APPLICANT: Evans, Sean
; APPLICANT: Shaked, Ze'ev
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS
; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,572
; FILING DATE: 07-June-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/445,307
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAIG, ANNE I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: 017.60S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6040
; INFORMATION FOR SEQ ID NO: 159:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 129 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-08-478-572-159

Query Match 12.6%; Score 100; DB 2; Length 129;
Best Local Similarity 94.1%; Pred. No. 0.00026;
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 52 GCHGSEPCIIHRGKPF 68
Db 20 GCHGSEPCIIHRGKPF 36
```



```

; TITLE OF INVENTION:  T CELL EPITOPES OF THE MAJOR ALLERGENS
; TITLE OF INVENTION:  FROM DERMATOPHAGOIDES (HOUSE DUST MITE)
; NUMBER OF SEQUENCES:  207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE:  IMMULOGIC PHARMACEUTICAL CORPORATION
; STREET:  610 LINCOLN STREET
; CITY:  WALTHAM
; STATE:  MA
; COUNTRY:  USA
; ZIP:  02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE:  Floppy disk
; COMPUTER:  IBM PC compatible
; OPERATING SYSTEM:  PC-DOS/MS-DOS
; SOFTWARE:  ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER:  US/08/484,296
; FILING DATE:
; CLASSIFICATION:  435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:  08/445,307
; FILING DATE:  07 June 1995
; ATTORNEY/AGENT INFORMATION:
; NAME:  CRAIG, ANNE I.
; REGISTRATION NUMBER:  32,976
; REFERENCE/DOCKET NUMBER:  017.6US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:  (617) 466-6000
; TELEFAX:  (617) 466-6040
; INFORMATION FOR SEQ ID NO: 159:
; SEQUENCE CHARACTERISTICS:
; LENGTH:  129 amino acids
; TYPE:  amino acid
; TOPOLOGY:  linear
; MOLECULE TYPE:  peptide
; FRAGMENT TYPE:  internal
; US-08-484-296-159

Query Match          12.6%;   Score 100;   DB 4;   Length 129;
Best Local Similarity 94.1%;   Pred. No. 0.00026;
Matches 16; Conservative 1; Mismatches 0; Indels

QY  52  GCHGSEPCIHRGKPF5 68
      | | | | | | | | | | | | | |
Db   20  GCHGSEPCIHRGKPF 36

Search completed: December  4, 2002, 12:55:07
Job time : 22.8333 secs

```

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 12:49:07 ; Search time 7.83333 Seconds  
(without alignments)  
490.899 Million cell updates/sec

Title: US-09-362-731a-4  
Perfect score: 203  
Sequence: 1 PKYVKQNTLKLATGCKGPKYVKQNTLKLATGCKGKGVIIIGIK 40

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR\_73:\*

2: pirl:\*

3: pirl3:\*

4: pirl4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85.5	42.1	565	1 HMIV65	hemagglutinin prec
2	85.5	42.1	565	1 HMIV66	hemagglutinin prec
3	85.5	42.1	565	1 HMIV68	hemagglutinin prec
4	82.5	40.6	565	1 HMIV64	hemagglutinin prec
5	82.5	40.6	565	1 HMIV69	hemagglutinin prec
6	82.5	40.6	565	1 HMIV67	hemagglutinin prec
7	82.5	40.6	565	1 HMIV6E	hemagglutinin prec
8	82.5	40.6	565	2 S33703	hemagglutinin - in
9	81.5	40.1	550	1 HMIV53	hemagglutinin prec
10	81.5	40.1	550	1 HMIV77	hemagglutinin prec
11	81.5	40.1	550	1 HMIV80	hemagglutinin prec
12	81.5	40.1	550	1 HMIV33	hemagglutinin prec
13	81.5	40.1	550	1 HMIV89	hemagglutinin prec
14	81.5	40.1	550	1 HMIV98	hemagglutinin prec
15	81.5	40.1	550	1 HMIV15	hemagglutinin prec
16	81.5	40.1	550	2 JQ1156	hemagglutinin prec
17	81.5	40.1	565	1 HMIV1	hemagglutinin prec
18	81.5	40.1	565	1 HMIV2	hemagglutinin prec
19	81.5	40.1	565	1 HMIV3	hemagglutinin prec
20	81.5	40.1	565	1 HMIV7	hemagglutinin prec
21	81.5	40.1	566	1 HMIVH	hemagglutinin prec
22	81.5	40.1	566	1 HMIVHA	hemagglutinin prec
23	81.5	40.1	566	1 HMIV6	hemagglutinin prec
24	81.5	40.1	566	1 HMIV6	hemagglutinin prec
25	81.5	40.1	566	1 HMIVDU	hemagglutinin prec
26	81.5	40.1	567	1 HMIVV	hemagglutinin prec
27	78.5	38.7	537	2 S52188	hemagglutinin - in
28	78.5	38.7	550	1 HMIV86	hemagglutinin prec
29	77.5	38.2	550	1 HMIV52	hemagglutinin prec

30	77.5	38.2	550	2 JQ1153	hemagglutinin prec
31	75.5	37.2	330	2 JQ2374	hemagglutinin - in
32	75.5	37.2	330	2 JQ2375	hemagglutinin - in
33	75.5	37.2	331	2 JQ2377	hemagglutinin - in
34	75.5	37.2	331	2 JQ2378	hemagglutinin - in
35	75.5	37.2	347	2 S52173	hemagglutinin - in
36	75.5	37.2	347	2 S52174	hemagglutinin - in
37	75.5	37.2	347	2 S52175	hemagglutinin - in
38	75.5	37.2	347	2 S52176	hemagglutinin - in
39	75.5	37.2	347	2 S52178	hemagglutinin - in
40	75.5	37.2	347	2 S52180	hemagglutinin - in
41	75.5	37.2	347	2 S52182	hemagglutinin - in
42	75.5	37.2	347	2 S52183	hemagglutinin - in
43	75.5	37.2	347	2 S52185	hemagglutinin - in
44	75.5	37.2	347	2 S52186	hemagglutinin - in
45	75.5	37.2	347	2 S52187	hemagglutinin - in

#### ALIGNMENTS

##### RESULT 1

HMIV65

hemagglutinin precursor - influenza A virus (strain A/equine/New Market/76[H3N8])  
N:Contains: hemagglutinin HA1; hemagglutinin HA2

C:Species: Influenza A virus

C>Date: 30-Sep-1990 #sequence\_revision 30-Sep-1990 #text\_change 16-Jul-1999

C:Accession: E34064

R:Kawaoka, Y.; Bean, W.J.; Webster, R.G.

Virology 169, 283-292, 1989

A:Title: Evolution of the hemagglutinin of equine H3 influenza viruses.

A:Reference number: A34064; MUID:89204899; PMID:2705299

A:Accession: E34064

A:Molecule type: genomic RNA

A:Residues: 1-565 <KAW>

A:Cross-references: GB:M24722; GB:J04336; NID:g324010; PIDN:AAA43107.1; PID:g324011

C:Genetics:

A:Map position: segment 4

C:Superfamily: Influenza virus hemagglutinin

C:Keywords: glycoprotein; hemagglutinin; homotrimer; lipoprotein; thiolester bond

F:1-16/Domain: signal sequence #status predicted <SIG>

F:17-344/Product: hemagglutinin HA1 #status predicted <HA1>

F:345-565/Product: hemagglutinin HA2 #status predicted <HA2>

F:535-551/Domain: transmembrane #status predicted <TM1>

F:23-37,53,78,180,300,498/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:29-481,67-292,79-91,154-488,296-320/bisulfide bonds: #status predicted

F:554,561,564/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 42.1%; Score 85.5; DB 1; Length 565;

Best Local Similarity 63.3%; Pred. No. 0.011;

Matches 19; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

QY 6 QNTLKLATGCKGPKYVKQNTLKLATGCKGV 35

DB 310 QNVNKVTYG-KCPKYIKQNTLKLATGMNV 338

##### RESULT 2

HMIV66

hemagglutinin precursor - influenza A virus (strain A/equine/Fontainebleau/76[H3N8])  
N:Contains: hemagglutinin HA1; hemagglutinin HA2

C:Species: Influenza A virus

C>Date: 30-Sep-1990 #sequence\_revision 30-Sep-1990 #text\_change 16-Jul-1999

C:Accession: F34064

R:Kawaoka, Y.; Bean, W.J.; Webster, R.G.

Virology 169, 283-292, 1989

A:Title: Evolution of the hemagglutinin of equine H3 influenza viruses.

A:Reference number: A34064; MUID:89204899; PMID:2705299

A:Accession: F34064

A:Molecule type: genomic RNA

A:Residues: 1-565 <KAW>

A:Cross-references: GB:M24723; GB:J04336; NID:g323998; PIDN:AAA43101.1; PID:g323999

C:Genetics:

A:Map position: segment 4  
C:Superfamily: influenza virus hemagglutinin  
C:Keywords: glycoprotein; hemagglutinin; homotrimer; lipoprotein; thiolester bond  
F:1-16/Domain: signal sequence #status predicted <SIG>  
F:17-344/Product: hemagglutinin HA1 #status predicted <HA1>  
F:345-565/Product: hemagglutinin HA2 #status predicted <HA2>  
F:535-551/Domain: transmembrane #status predicted <TM1>  
F:18,23,37,53,78,180,300,498/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:29-481,67-292,79-91,154-488,296-320/Disulfide bonds: #status predicted  
F:554,561,564/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 42.18; Score 85.5; DB 1; Length 565;  
Best Local Similarity 63.3%; Pred. No. 0.011;  
Matches 19; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

Qy 6 QNTLKLATGKGGPKYVKONTLKLATGKGGV 35  
|| | : | | | : | | | | | | | | | : |  
Db 310 QNVNKTYYG-KCPKYIKONTLKLATGMNRV 338

RESULT 3  
HMIVE8  
hemagglutinin precursor - influenza A virus (strain A/equine/Santiago/1/85[H3N8])  
N:Contains: hemagglutinin HA1; hemagglutinin HA2  
C:Species: influenza A virus  
C:Date: 30-Sep-1990 #sequence\_revision 30-Sep-1990 #text\_change 16-Jul-1999  
C:Accession: H34064  
R:Kawaoka, Y.; Bean, W.J.; Webster, R.G.  
Virology 169, 283-292, 1989  
A:Title: Evolution of the hemagglutinin of equine H3 influenza viruses.  
A:Reference number: A34064; MUID:89204899; PMID:2705299  
A:Accession: H34064  
A:Molecule type: genomic RNA  
A:Residues: 1-565 <KAW>  
A:Cross-references: GB:M24725; GB:J04336; NID:g324016; PIDN:AAA43110.1; PID:g324017  
C:Genetics:  
A:Map position: segment 4  
C:Superfamily: influenza virus hemagglutinin  
C:Keywords: glycoprotein; hemagglutinin; homotrimer; lipoprotein; thiolester bond  
F:1-16/Domain: signal sequence #status predicted <SIG>  
F:17-344/Product: hemagglutinin HA1 #status predicted <HA1>  
F:345-565/Product: hemagglutinin HA2 #status predicted <HA2>  
F:535-551/Domain: transmembrane #status predicted <TM1>  
F:18,23,37,53,78,180,300,498/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:29-481,67-292,79-91,154-488,296-320/Disulfide bonds: #status predicted  
F:554,561,564/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 42.18; Score 85.5; DB 1; Length 565;  
Best Local Similarity 63.3%; Pred. No. 0.011;  
Matches 19; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

Qy 6 QNTLKLATGKGGPKYVKONTLKLATGKGGV 35  
|| | : | | | : | | | | | | | | | : |  
Db 310 QNVNKTYYG-KCPKYIKONTLKLATGMNRV 338

RESULT 4  
HMIVE4  
hemagglutinin precursor - influenza A virus (strain A/equine/Algiers/72[H3N8])  
N:Contains: hemagglutinin HA1; hemagglutinin HA2  
C:Species: influenza A virus  
C:Date: 30-Sep-1990 #sequence\_revision 30-Sep-1990 #text\_change 16-Jul-1999  
C:Accession: D34064  
R:Kawaoka, Y.; Bean, W.J.; Webster, R.G.  
Virology 169, 283-292, 1989  
A:Title: Evolution of the hemagglutinin of equine H3 influenza viruses.  
A:Reference number: A34064; MUID:89204899; PMID:2705299  
A:Accession: D34064  
A:Molecule type: genomic RNA  
A:Residues: 1-565 <KAW>  
A:Cross-references: GB:M24721; GB:J04336; NID:g323996; PIDN:AAA43100.1; PID:g323997  
C:Genetics:  
A:Map position: segment 4

C:Superfamily: influenza virus hemagglutinin  
C:Keywords: glycoprotein; hemagglutinin; homotrimer; lipoprotein; thiolester bond  
F:1-16/Domain: signal sequence #status predicted <SIG>  
F:17-344/Product: hemagglutinin HA1 #status predicted <HA1>  
F:345-565/Product: hemagglutinin HA2 #status predicted <HA2>  
F:535-551/Domain: transmembrane #status predicted <TM1>  
F:18,23,37,53,78,180,300,498/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:29-481,67-292,79-91,154-488,296-320/Disulfide bonds: #status predicted  
F:554,561,564/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 40.6%; Score 82.5; DB 1; Length 565;  
Best Local Similarity 63.3%; Pred. No. 0.025;  
Matches 19; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

Qy 6 QNTLKLATGKGGPKYVKONTLKLATGKGGV 35  
|| | : | | | : | | | | | | | | | : |  
Db 310 QNVNKTYYG-KCPKYIKONTLKLATGMNRV 338

RESULT 5  
HMIVE9  
hemagglutinin precursor - influenza A virus (strain A/equine/Tennessee/5/85[H3N8])  
N:Contains: hemagglutinin HA1; hemagglutinin HA2  
C:Species: influenza A virus  
C:Date: 30-Sep-1990 #sequence\_revision 30-Sep-1990 #text\_change 16-Jul-1999  
C:Accession: I34064  
R:Kawaoka, Y.; Bean, W.J.; Webster, R.G.  
Virology 169, 283-292, 1989  
A:Title: Evolution of the hemagglutinin of equine H3 influenza viruses.  
A:Reference number: A34064; MUID:89204899; PMID:2705299  
A:Accession: I34064  
A:Molecule type: genomic RNA  
A:Residues: 1-565 <KAW>  
A:Cross-references: GB:M24726; GB:J04336; NID:g324020; PIDN:AAA43112.1; PID:g324021  
C:Genetics:  
A:Map position: segment 4  
C:Superfamily: influenza virus hemagglutinin  
C:Keywords: glycoprotein; hemagglutinin; homotrimer; lipoprotein; thiolester bond  
F:1-16/Domain: signal sequence #status predicted <SIG>  
F:17-344/Product: hemagglutinin HA1 #status predicted <HA1>  
F:345-565/Product: hemagglutinin HA2 #status predicted <HA2>  
F:535-551/Domain: transmembrane #status predicted <TM1>  
F:18,23,37,53,78,180,300,498/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:29-481,67-292,79-91,154-488,296-320/Disulfide bonds: #status predicted  
F:554,561,564/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 40.6%; Score 82.5; DB 1; Length 565;  
Best Local Similarity 60.0%; Pred. No. 0.025;  
Matches 18; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

Qy 6 QNTLKLATGKGGPKYVKONTLKLATGKGGV 35  
|| | : | | | : | | | | | | | | | : |  
Db 310 QNVNKTYYG-KCPKYIKONTLKLATGMNRV 338

RESULT 6  
HMIVET  
hemagglutinin precursor - influenza A virus (strain A/equine/Kentucky/2/86[H3N8])  
N:Contains: hemagglutinin HA1; hemagglutinin HA2  
C:Species: influenza A virus  
C:Date: 30-Sep-1990 #sequence\_revision 30-Sep-1990 #text\_change 16-Jul-1999  
C:Accession: A34065  
R:Kawaoka, Y.; Bean, W.J.; Webster, R.G.  
Virology 169, 283-292, 1989  
A:Title: Evolution of the hemagglutinin of equine H3 influenza viruses.  
A:Reference number: A34064; MUID:89204899; PMID:2705299  
A:Accession: A34065  
A:Molecule type: genomic RNA  
A:Residues: 1-565 <KAW>  
A:Cross-references: GB:M24727; GB:J04336; NID:g324000; PIDN:AAA43102.1; PID:g324001  
C:Genetics:  
A:Map position: segment 4  
C:Superfamily: influenza virus hemagglutinin

Best Local Similarity 60.0%; Pred. No. 0.025;  
Matches 18; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 6 QNTLKLATGKGGPKYVVKONTLKLATGKKGV 35  
|| | : | ||| : ||||| : |  
DB 310 QNVNKVTYG-KCPKYIIRONTLKLATGMNRV 338

RESULT 9  
HMIIV53  
hemagglutinin precursor - influenza A virus (strain A/swine/81/78) (fragment)  
C:Species: Influenza A virus  
C:Date: 31-Mar-1989 #sequence-revision 31-Mar-1989 #text\_change 18-Sep-1998  
C:Accession: B29971  
R:Kida, H.; Shortridge, K.F.; Webster, R.G.  
Virology 162, 160-166, 1988  
A:Title: Origin of the hemagglutinin gene of H3N2 influenza viruses from pigs in China  
A:Reference number: A94370; MUID:88101364; PMID:3336940  
A:Accession: B29971  
A:Molecule type: genomic RNA  
A:Residues: 1-550 <KID>  
A:Cross-references: GB:M19057; NID:g324210  
A>Note: the sequence in GenBank entry FLAHABP, release 106, (PID:g324211) differs from  
C:Genetics:  
A:Map position: segment 4  
C:Superfamily: influenza virus hemagglutinin  
C:Keywords: glycoprotein; hemagglutinin; homotrimer; lipoprotein; thiolester bond  
F:1-328/Product: hemagglutinin HA1 #status predicted <HA1>  
F:300-550/Product: hemagglutinin HA2 #status predicted <HA2>  
F:520-536/Domain: transmembrane #status predicted <TM1>  
F:8:22,38,165,285,483/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:14-466,52-277,64-76,139-473,281-305/Disulfide bonds: #status predicted  
F:539,546,549/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 40.1%; Score 81.5; DB 1; Length 550;  
Best Local Similarity 63.3%; Pred. No. 0.033;  
Matches 19; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

QY 6 QNTLKLATGKGGPKYVVKONTLKLATGKKGV 35  
|| | : | ||| : ||||| : |  
DB 295 QNVNKVTYG-ACPYYVVKONTLKLATGMNRV 323

RESULT 10  
HMIIV77  
hemagglutinin precursor - influenza A virus (strain A/duck/Hokkaido/5/77) (fragment)  
N:Contains: hemagglutinin HA1; hemagglutinin HA2  
C:Species: Influenza A virus  
C:Date: 30-Jun-1989 #sequence-revision 30-Jun-1989 #text\_change 16-Jul-1999  
C:Accession: A27813  
R:Kida, H.; Kawacka, Y.; Naeye, C.W.; Webster, R.G.  
Virology 159, 109-119, 1987  
A:Title: Antigenic and genetic conservation of H3 influenza virus in wild ducks.  
A:Reference number: A94363; MUID:87265459; PMID:2440178  
A:Accession: A27813  
A:Molecule type: genomic RNA  
A:Residues: 1-550 <KID>  
A:Cross-references: GB:M16737; NID:g324081; PIDN:AAA43143.1; PID:g324082  
C:Genetics:  
A:Map position: segment 4  
C:Superfamily: influenza virus hemagglutinin  
C:Keywords: glycoprotein; hemagglutinin; homotrimer; lipoprotein; thiolester bond  
F:1-328/Product: hemagglutinin HA1 #status predicted <HA1>  
F:330-550/Product: hemagglutinin HA2 #status predicted <HA2>  
F:520-536/Domain: transmembrane #status predicted <TM1>  
F:8:22,38,165,285,483/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:14-466,52-277,64-76,139-473,281-305/Disulfide bonds: #status predicted  
F:539,546,549/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 40.1%; Score 81.5; DB 1; Length 550;  
Best Local Similarity 63.3%; Pred. No. 0.033;  
Matches 19; Conservative 2; Mismatches 8; Indels 1; Gaps 1;



```

RESULT 15
HMIIV5.
hemagglutinin precursor - influenza A virus (strain A/duck/Hokkaido/10/85) (fragment)
N;Contnns: hemagglutinin HA1; hemagglutinin HA2
C;Species: Influenza A virus
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 16-Jul-1999
C;Accession: G27813
R;Kida, H.; Kawaoka, Y.; Naeye, C.W.; Webster, R.G.
Virology 159, 109-119, 1987
A;Title: Antigenic and genetic conservation of H3 influenza virus in wild ducks.
A;Reference number: A94363; MUID:87265458; PMID:2440178
A;Accession: G27813
A:Molecule type: genomic RNA
A;Residues: 1-550 <KID>
A;Cross-references: GB:M16743; NID:g324093; PIDN:AAA43149.1; PID:g324094
C;Genetics:
A;Map position: segment 4
C;Superfamily: Influenza virus hemagglutinin
C;Keywords: glycoprotein; hemagglutinin; homotrimer; lipoprotein; thiolester bond
F;1-328/Product: hemagglutinin HA1 #status predicted <HA1>
F;330-550/Product: hemagglutinin HA2 #status predicted <HA2>
F;520-536/Domain: transmembrane #status predicted <TM1>
F;8,22,38,165,285,483/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;14-466,52-277,64-76,139-473,281-305/Disulfide bonds: #status predicted
F;539,546,549/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match          40.1%; Score 81.5; DB 1; Length 550;
Best Local Similarity 63.3%; Pred. No. 0 033;
Matches 19; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

Qy      6 QNTLKLATGKGPKYVKONTLKLATGKKGV 35
Db      295 QNVNKITY-ACPKYVKONTLKLATGMNV 323

```

Search completed: December 4, 2002, 12:54:23  
Job time : 8.83333 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 12:49:06 ; Search time 4.3333 Seconds  
(without alignments)  
382.858 Million cell updates/sec

Title: US-09-362-731A-4  
Perfect score: 203  
Sequence: 1 PKYVKQNTLKLATGKGPYVKQNTLKLATGKKGVIIGIK 40

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	85.5	42.1	565	1 HEMA_IAHFO	P16995 influenza a
2	85.5	42.1	565	1 HEMA_IAHNM	P16997 influenza a
3	85.5	42.1	565	1 HEMA_IAHSA	P16999 influenza a
4	83.5	41.1	566	1 HEMA_IADDO	P19106 influenza a
5	82.5	40.6	565	1 HEMA_IADAL	P16994 influenza a
6	82.5	40.6	565	1 HEMA_IADH6	P19699 influenza a
7	82.5	40.6	565	1 HEMA_IADH7	P16996 influenza a
8	82.5	40.6	565	1 HEMA_IADSU	O08011 influenza a
9	82.5	40.6	565	1 HEMA_IADTE	P17001 influenza a
10	81.5	40.1	328	1 HEMA_IAEN6	P04664 influenza a
11	81.5	40.1	328	1 HEMA_IADU7	P04663 influenza a
12	81.5	40.1	550	1 HEMA_IADAN	P03441 influenza a
13	81.5	40.1	550	1 HEMA_IADH1	P12582 influenza a
14	81.5	40.1	550	1 HEMA_IADH2	P12583 influenza a
15	81.5	40.1	550	1 HEMA_IADH3	P12584 influenza a
16	81.5	40.1	550	1 HEMA_IADH4	P12585 influenza a
17	81.5	40.1	550	1 HEMA_IADH6	P12587 influenza a
18	81.5	40.1	550	1 HEMA_IADH7	P12588 influenza a
19	81.5	40.1	550	1 HEMA_IADHM	P43259 influenza a
20	81.5	40.1	550	1 HEMA_IADH2	P11133 influenza a
21	81.5	40.1	565	1 HEMA_IADH1	P15658 influenza a
22	81.5	40.1	565	1 HEMA_IADH1	P16998 influenza a
23	81.5	40.1	565	1 HEMA_IADH1	P17000 influenza a
24	81.5	40.1	565	1 HEMA_IADH1	P17002 influenza a
25	81.5	40.1	566	1 HEMA_IADH1	P03437 influenza a
26	81.5	40.1	566	1 HEMA_IADU3	P03442 influenza a
27	81.5	40.1	566	1 HEMA_IADH1	P03440 influenza a
28	81.5	40.1	566	1 HEMA_IADH1	P03439 influenza a
29	81.5	40.1	566	1 HEMA_IADH1	P03436 influenza a
30	81.5	40.1	566	1 HEMA_IADH1	P03435 influenza a
31	81.5	40.1	567	1 HEMA_IADH1	P26139 influenza a
32	80.5	39.7	566	1 HEMA_IADH1	P03435 influenza a
33	78.5	38.7	550	1 HEMA_IADH1	P26141 influenza a
					P12589 influenza a

34	77.5	38.2	550	1 HEMA_IADHK	P43257 influenza a
35	77.5	38.2	550	1 HEMA_IADH3	P11134 influenza a
36	76.5	37.7	566	1 HEMA_IADH2	P26135 influenza a
37	75.5	37.2	550	1 HEMA_IADHL	P43258 influenza a
38	75.5	37.2	550	1 HEMA_IADHK	P43260 influenza a
39	75.5	37.2	564	1 HEMA_IADUD	P19694 influenza a
40	75.5	37.2	566	1 HEMA_IADH3	P26134 influenza a
41	75.5	37.2	566	1 HEMA_IADAO	P26138 influenza a
42	73.5	36.2	566	1 HEMA_IADH1	P03449 influenza a
43	70.5	34.7	564	1 HEMA_IADH2	P19696 influenza a
44	70.5	34.7	564	1 HEMA_IADH1	P19698 influenza a
45	69.5	34.2	564	1 HEMA_IADH1	P19695 influenza a

#### ALIGNMENTS

RESULT 1

HEMA\_IAHFO

ID HEMA\_IAHFO STANDARD; PRT; 565 AA.

AC P16995; Q83993; Q83992;

DT 01-AUG-1990 (Rel. 15, Created)

DT 01-FEB-1991 (Rel. 17, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hemagglutinin precursor [Contains: Hemagglutinin HA1 chain;

DE Hemagglutinin HA2 chain].

GN HA.

OS Influenza A virus (strain A/Equine/Fontainebleau/76) (Influenza A

OC virus (strain A/Equine/France/1/76)).

OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;

OC Influenza A viruses; Influenzavirus A.

OC NCBI\_TaxID=11399;

RP [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=89204899; PubMed=2705299;

RA Kawaoka Y., Bean W.J., Webster R.G.;

RT "Evolution of the hemagglutinin of equine H3 influenza viruses.;"

RL Virology 169:283-292(1989).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=92114135; PubMed=1731092;

RA Bean W.J., Schell M., Katz J., Kawaoka Y., Naeve C., Gorman O.,

RT Webster R.G.;

RT "Evolution of the H3 influenza virus hemagglutinin from human and

nonhuman hosts.;"

RL J. Virol. 66:1129-1138(1992).

CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO

CELL RECEPTORS AND FOR INITIATING INFECTION.

CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS

(HA1 AND HA2) LINKED BY A DISULFIDE BOND.

CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; M24723; AAA43101.1; ALT\_SEQ.

CC EMBL; M73773; NOT\_ANNOTATED\_CDS.

DR PIR; F34064; HMIVP6.

DR HSSP; P03437; LHTW.

DR InterPro; IPR001364; Hemagglutn.

DR Pfam; PF00509; Hemagglutinin; 1.

```
FT CARBOHYD 37 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 53 78 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 78 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 180 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 300 498 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 498 565 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 13 15 WVY -> AVD (IN REF. 2).
FT CONFLICT 20 20 T -> I (IN REF. 2).
FT CONFLICT 150 150 R -> G (IN REF. 2).
FT CONFLICT 187 187 N -> D (IN REF. 2).
FT CONFLICT 242 242 S -> A (IN REF. 2).
FT CONFLICT 293 293 V -> W (IN REF. 2).
FT CONFLICT 479 479 N -> G (IN REF. 2).
FT CONFLICT 555 555 Q -> E (IN REF. 2).
SQ SEQUENCE 565 AA; 63686 MW; 1BB06B765982E87C CRC64;

Query Match 42.1%; Score 85.5; DB 1; Length 565;
Best Local Similarity 63.3%; Pred. No. 0.002;
Matches 19; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

Qy 6 QNTLKATGKKGPKYVKQNTLKATGKGV 35
|| | : | | | : | | | | | : |
Db 310 QNVNKVTYG-KCPKYIKQNTLKATGMRNV 338

RESULT 2
HEMA_IAHNM
ID HEMA_IAHNM STANDARD; PRT; 565 AA.
AC P16997; O83996;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hemagglutinin precursor [Contains: Hemagglutinin HA1 chain;
DE Hemagglutinin HA2 chain].
GN HA.
OS Influenza A virus (strain A/Equine/Santiago/1/85).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=11408;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89204899; PubMed=2705299;
RA Kawaoka Y., Bean W.J., Webster R.G.;
RT "Evolution of the hemagglutinin of equine H3 influenza viruses.";
RL Virology 169:283-292(1989).
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION.
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M24722; AAA43107.1; ALT_SEQ.
CC PIR; E34064; HMIVE5.
CC HSSP; P03437; IHTM.
CC InterPro; IPR001364; Hemagglutn.
CC Pfam; PF00509; Hemagglutinin; 1.
CC PRINTS; PR00329; HEMAGGLUTN12.
CC ProDom; PD000225; Hemagglutn; 1.
CC Envelope protein; Hemagglutinin; Glycoprotein; Signal.
CC SIGNAL 1 16
CC CHAIN 17 343 HEMAGGLUTININ HA1 CHAIN.
CC CHAIN 345 565 HEMAGGLUTININ HA2 CHAIN.
CC FT CARBOHYD 23 23 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 78 78 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 498 498 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SQ SEQUENCE 565 AA; 63665 MW; 399F4BF4BA231327 CRC64;

Query Match 42.1%; Score 85.5; DB 1; Length 565;
Best Local Similarity 63.3%; Pred. No. 0.002;
Matches 19; Conservative 3; Mismatches 7; Indels 1; Gaps 1;
```

```
FT CARBOHYD 78 78 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 498 498 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 565 AA; 63749 MW; 35B4D05C374FB215 CRC64;

Query Match 42.1%; Score 85.5; DB 1; Length 565;
Best Local Similarity 63.3%; Pred. No. 0.002;
Matches 19; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

Qy 6 QNTLKATGKKGPKYVKQNTLKATGKGV 35
|| | : | | | : | | | | | : |
Db 310 QNVNKVTYG-KCPKYIKQNTLKATGMRNV 338

RESULT 3
HEMA_IAHSA
ID HEMA_IAHSA STANDARD; PRT; 565 AA.
AC P16999; O84000; Q84001;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hemagglutinin precursor [Contains: Hemagglutinin HA1 chain;
DE Hemagglutinin HA2 chain].
GN HA.
OS Influenza A virus (strain A/Equine/Santiago/1/85).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=11414;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89204899; PubMed=2705299;
RA Kawaoka Y., Bean W.J., Webster R.G.;
RT "Evolution of the hemagglutinin of equine H3 influenza viruses.";
RL Virology 169:283-292(1989).
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION.
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M24725; AAA43110.1; ALT_SEQ.
CC PIR; H34064; HMIVE8.
CC HSSP; P03437; IHTM.
CC InterPro; IPR001364; Hemagglutn.
CC Pfam; PF00509; Hemagglutinin; 1.
CC PRINTS; PR00329; HEMAGGLUTN12.
CC ProDom; PD000225; Hemagglutn; 1.
CC Envelope protein; Hemagglutinin; Glycoprotein; Signal.
CC SIGNAL 1 16
CC CHAIN 17 343 HEMAGGLUTININ HA1 CHAIN.
CC CHAIN 345 565 HEMAGGLUTININ HA2 CHAIN.
CC FT CARBOHYD 23 23 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 78 78 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 498 498 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SQ SEQUENCE 565 AA; 63665 MW; 399F4BF4BA231327 CRC64;

Query Match 42.1%; Score 85.5; DB 1; Length 565;
Best Local Similarity 63.3%; Pred. No. 0.002;
Matches 19; Conservative 3; Mismatches 7; Indels 1; Gaps 1;
```





```

RN  SEQUENCE FROM N.A.
RX  MEDLINE-89204899; PubMed-2705299;
RA  Kawaoka Y., Bean W.J., Webster R.G.;
RT  "Evolution of the hemagglutinin of equine H3 influenza viruses.";
RL  Virology 169:283-292(1989).
CC  -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC  CELL RECEPTORS AND FOR INITIATING INFECTION.
CC  -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC  (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC  -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; M24727; -; NOT_ANNOTATED_CDS.
CC  PIR; A34065; HMIVET.
CC  HSSP; P03437; IHMT.
CC  InterPro; IPR001364; Hemagglutn.
CC  Pfam; PF00509; Hemagglutinin; 1.
CC  PRINTS; PR00329; HEMAGGLUTN12.
CC  ProDom; PD000225; Hemagglutn; 1.
CC  Envelope protein; Hemagglutinin; Glycoprotein; Signal.
FT  SIGNAL 1 16
FT  CHAIN 17 343 HEMAGGLUTININ HA1 CHAIN.
FT  CHAIN 345 565 HEMAGGLUTININ HA2 CHAIN.
FT  CARBOHYD 23 23 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 78 78 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 498 498 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ  SEQUENCE 565 AA; 63610 MW; 2038CCL1C6C9B88C5 CRC64;

Query Match 40.6%; Score 82.5; DB 1; Length 565;
Best Local Similarity 60.0%; Pred. No. 0.0048;
Matches 18; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 6 QNTLKATGKGPYKQNTLKATGKGV 35
    || | : | | | : | | | | | | | : |
DB 310 QNVNKVTYG-KCPKYIRQNTLKATGMNV 338

RESULT 7
HEMA_IAHK7
ID HEMA_IAHK7 STANDARD; PRT; 565 AA.
AC P16996; O83995; O83994;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hemagglutinin precursor [Contains: Hemagglutinin HA1 chain;
DE Hemagglutinin HA2 chain].
GN HA.
OS Influenza A virus (strain A/Kentucky/1/87).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=11404;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-89204899; PubMed-2705299;
RA Kawaoka Y., Bean W.J., Webster R.G.;
RT "Evolution of the hemagglutinin of equine H3 influenza viruses.";
RL Virology 169:283-292(1989).
CC  -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC  CELL RECEPTORS AND FOR INITIATING INFECTION.
CC  -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC  (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC  -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; M24727; -; NOT_ANNOTATED_CDS.
CC  PIR; A34065; HMIVET.
CC  HSSP; P03437; IHMT.
CC  InterPro; IPR001364; Hemagglutn.
CC  Pfam; PF00509; Hemagglutinin; 1.
CC  PRINTS; PR00329; HEMAGGLUTN12.
CC  ProDom; PD000225; Hemagglutn; 1.
CC  Envelope protein; Hemagglutinin; Glycoprotein; Signal.
FT  SIGNAL 1 16
FT  CHAIN 17 343 HEMAGGLUTININ HA1 CHAIN.
FT  CHAIN 345 565 HEMAGGLUTININ HA2 CHAIN.
FT  CARBOHYD 23 23 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 78 78 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 498 498 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ  SEQUENCE 565 AA; 63610 MW; 2038CCL1C6C9B88C5 CRC64;

Query Match 40.6%; Score 82.5; DB 1; Length 565;
Best Local Similarity 60.0%; Pred. No. 0.0048;
Matches 18; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 6 QNTLKATGKGPYKQNTLKATGKGV 35
    || | : | | | : | | | | | | | : |
DB 310 QNVNKVTYG-KCPKYIRQNTLKATGMNV 338

RESULT 8
HEMA_IAHSU
ID HEMA_IAHSU STANDARD; PRT; 565 AA.
AC O08011;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hemagglutinin precursor [Contains: Hemagglutinin HA1 chain;
DE Hemagglutinin HA2 chain].
GN HA.
OS Influenza A virus (strain A/Equine/Suffolk/89).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=45413;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-93277383; PubMed-8503788;
RX Binns M.M., Daly J.M., Chirnside E.D., Mumford J.A., Wood J.M.,
RX Richards C.M., Daniels R.S.;
RT "Genetic and antigenic analysis of an equine influenza H 3 isolate
RT from the 1989 epidemic.";
RL Arch. Virol. 130:33-44(1993).
CC  -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC  CELL RECEPTORS AND FOR INITIATING INFECTION.
CC  -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC  (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC  -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
```

```

CC  (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC  -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; M24728; AAA43103.1; ALT_SEQ.
CC  PIR; B34065; HMIVEE.
CC  HSSP; P03437; IHMT.
CC  InterPro; IPR001364; Hemagglutn.
CC  Pfam; PF00509; Hemagglutinin; 1.
CC  PRINTS; PR00329; HEMAGGLUTN12.
CC  ProDom; PD000225; Hemagglutn; 1.
CC  Envelope protein; Hemagglutinin; Glycoprotein; Signal.
FT  SIGNAL 1 16
FT  CHAIN 17 343 HEMAGGLUTININ HA1 CHAIN.
FT  CHAIN 345 565 HEMAGGLUTININ HA2 CHAIN.
FT  CARBOHYD 23 23 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 78 78 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 498 498 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ  SEQUENCE 565 AA; 63702 MW; 93963AF456486787 CRC64;

Query Match 40.6%; Score 82.5; DB 1; Length 565;
Best Local Similarity 60.0%; Pred. No. 0.0048;
Matches 18; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 6 QNTLKATGKGPYKQNTLKATGKGV 35
    || | : | | | : | | | | | | | : |
DB 310 QNVNKVTYG-KCPKYIRQNTLKATGMNV 338

RESULT 8
HEMA_IAHSU
ID HEMA_IAHSU STANDARD; PRT; 565 AA.
AC O08011;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hemagglutinin precursor [Contains: Hemagglutinin HA1 chain;
DE Hemagglutinin HA2 chain].
GN HA.
OS Influenza A virus (strain A/Equine/Suffolk/89).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=45413;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-93277383; PubMed-8503788;
RX Binns M.M., Daly J.M., Chirnside E.D., Mumford J.A., Wood J.M.,
RX Richards C.M., Daniels R.S.;
RT "Genetic and antigenic analysis of an equine influenza H 3 isolate
RT from the 1989 epidemic.";
RL Arch. Virol. 130:33-44(1993).
CC  -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC  CELL RECEPTORS AND FOR INITIATING INFECTION.
CC  -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC  (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC  -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
```

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```
CC EMBL; X68437; CAA48482.1; -
DR HSSP; P03437; IHTM.
DR InterPro; IPR001364; Hemagglutn.
DR Pfam; PF00509; Hemagglutinin; 1.
DR PRINTS; PR00329; Hemagglutn12.
DR ProDom; PD000225; Hemagglutn; 1.
KW Envelope protein; Hemagglutinin; Glycoprotein; Signal.
FT SIGNAL 1 16 BY SIMILARITY.
FT CHAIN 17 343 HEMAGGLUTININ HA1 CHAIN (BY SIMILARITY).
FT CHAIN 345 565 HEMAGGLUTININ HA2 CHAIN (BY SIMILARITY).
FT CARBOHYD 23 23 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 37 37 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 53 53 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 68 68 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 78 78 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 180 180 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 498 498 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 565 AA; 63699 MW; C7A4E3B54B87D1A1 CRC64;

Query Match 40.6%; Score 82.5; DB 1; Length 565;
Best Local Similarity 60.0%; Pred. No. 0.0048;
Matches 18; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

Qy 6 QNTLKATGKKGPKYVKONTLKLATGKGV 35
|| : | | | | | | | | | | | | | |
Db 310 QNVNKVTYG-KCPKYIRONTLKLATGMNV 338

RESULT 9
HEMA_IAHTE
ID HEMA_IAHTE STANDARD; PRT; 565 AA.
AC P17001; Q84004; Q84005;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hemagglutinin precursor [Contains: Hemagglutinin HA1 chain; Hemagglutinin HA2 chain].
GN HA.
OS Influenza A virus (strain A/Equine/Tennessee/5/86).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=11417;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89204899; PubMed=2705299;
RA Kawaoka Y., Bean W.J., Webster R.G.;
RT "Evolution of the hemagglutinin of equine H3 influenza viruses.";
RL Virology 169;283-292(1989).
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO CELL RECEPTORS AND FOR INITIATING INFECTION.
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
```

```
DR ProDom; PD000225; Hemagglutn; 1.
KW Envelope protein; Hemagglutinin; Glycoprotein; Signal.
FT SIGNAL 1 16
FT CHAIN 17 343 HEMAGGLUTININ HA1 CHAIN.
FT CHAIN 345 565 HEMAGGLUTININ HA2 CHAIN.
FT CARBOHYD 23 23 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 37 37 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 53 53 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 68 68 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 78 78 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 180 180 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 498 498 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 565 AA; 63688 MW; 1FE4485F0E7AC2C4 CRC64;

Query Match 40.6%; Score 82.5; DB 1; Length 565;
Best Local Similarity 60.0%; Pred. No. 0.0048;
Matches 18; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

Qy 6 QNTLKATGKKGPKYVKONTLKLATGKGV 35
|| : | | | | | | | | | | | | | |
Db 310 QNVNKVTYG-KCPKYIRONTLKLATGMNV 338

RESULT 10
HEMA_IAEN6
ID HEMA_IAEN6 STANDARD; PRT; 328 AA.
AC P04664;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hemagglutinin [Contains: Hemagglutinin HA1 chain] (Fragment).
GN HA.
OS Influenza A virus (strain A/England/878/69).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=11377;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81194918; PubMed=6164798;
RA Sleight M.J., Both G.W., Underwood P.A., Bender V.J.;
RT "Antigenic drift in the hemagglutinin of the Hong Kong influenza subtype: correlation of amino acid changes with alterations in viral antigenicity.";
RT
RT Influenza A virus (strain A/England/878/69).
RL J. Virol. 37;845-853(1981).
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO CELL RECEPTORS AND FOR INITIATING INFECTION.
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
```

```
CC EMBL; K03335; AAA43184.1; -
DR HSSP; P03437; IHGE.
DR InterPro; IPR001364; Hemagglutn.
DR Pfam; PF00509; Hemagglutinin; 1.
DR ProDom; PD000225; Hemagglutn; 1.
KW Envelope protein; Hemagglutinin; Glycoprotein.
FT NON_TER 1 1
FT CHAIN 1 328 HEMAGGLUTININ HA1 CHAIN.
FT CARBOHYD 8 8 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 22 22 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 38 38 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 63 63 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 165 165 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 285 285 N-LINKED (GLCNAC. .) (POTENTIAL).
```



```

GN HA.
OS Influenza A virus (strain A/Duck/Hokkaido/5/77).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
CC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=11357;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87265458; PubMed=2440178;
RA Kida H., Kawaoka Y., Naeve C.W., Webster R.G.;
RT "Antigenic and genetic conservation of H3 influenza virus in wild ducks.";
RL Virology 159:109-119(1987).
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO CELL RECEPTORS AND FOR INITIATING INFECTION.
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M16737; AAA43143.1; -.
DR PIR; A27813; HMIV77.
DR HSSP; P03437; 3HMG.
DR InterPro; IPR001364; Hemagglutn.
DR Pfam; PF00509; Hemagglutinin; 1.
DR ProDom; PD000225; Hemagglutn; 1.
DR Envelope protein; Hemagglutinin; Glycoprotein.
FT NON_TER 1 1
FT CHAIN 1 328 HEMAGGLUTININ HA1 CHAIN.
FT CHAIN 330 550 HEMAGGLUTININ HA2 CHAIN.
FT CARBOHYD 8 8 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 285 285 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 137 137 K -> N (IN PIR DATA BANK).
SQ SEQUENCE 550 AA; 61705 MW; 7E7ACEF716FC969A CRC64;

Qy 6 QNTLKLATGKGGPKYKQNTLKLATGKGGV 35
|| | : | ||||| ||||| : |
Db 295 QNVNKITYG-ACPKYKQNTLKLATGMNV 323

RESULT 14
HEMA_IADH2
ID HEMA_IADH2 STANDARD; PRT; 550 AA.
AC P12583; Q84011;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hemagglutinin precursor [Contains: Hemagglutinin HA1 chain; Hemagglutinin HA2 chain] (Fragment).
GN HA.
OS Influenza A virus (strain A/Duck/Hokkaido/8/80).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
CC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=11358;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87265456; PubMed=2440178;
RA Kida H., Kawaoka Y., Naeve C.W., Webster R.G.;
RT "Antigenic and genetic conservation of H3 influenza virus in wild ducks.";
RL Virology 159:109-119(1987).
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO CELL RECEPTORS AND FOR INITIATING INFECTION.
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M16737; AAA43143.1; -.
DR PIR; A27813; HMIV77.
DR HSSP; P03437; 3HMG.
DR InterPro; IPR001364; Hemagglutn.
DR Pfam; PF00509; Hemagglutinin; 1.
DR ProDom; PD000225; Hemagglutn; 1.
DR Envelope protein; Hemagglutinin; Glycoprotein.
FT NON_TER 1 1
FT CHAIN 1 328 HEMAGGLUTININ HA1 CHAIN.
FT CHAIN 330 550 HEMAGGLUTININ HA2 CHAIN.
FT CARBOHYD 8 8 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 285 285 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 137 137 K -> N (IN PIR DATA BANK).
SQ SEQUENCE 550 AA; 61705 MW; 7E7ACEF716FC969A CRC64;

Qy 6 QNTLKLATGKGGPKYKQNTLKLATGKGGV 35
|| | : | ||||| ||||| : |
Db 295 QNVNKITYG-ACPKYKQNTLKLATGMNV 323

RESULT 15
HEMA_IADH3
ID HEMA_IADH3 STANDARD; PRT; 550 AA.
AC P12584; Q84012; Q89793;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hemagglutinin precursor [Contains: Hemagglutinin HA1 chain; Hemagglutinin HA2 chain] (Fragment).
GN HA.
OS Influenza A virus (strain A/Duck/Hokkaido/33/80).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
CC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=11359;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87265458; PubMed=2440178;
RA Kida H., Kawaoka Y., Naeve C.W., Webster R.G.;
RT "Antigenic and genetic conservation of H3 influenza virus in wild ducks.";
RL Virology 159:109-119(1987).
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO CELL RECEPTORS AND FOR INITIATING INFECTION.
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M16738; AAA43144.1; -.
DR PIR; B27813; HMIV80.
DR HSSP; P03437; 2VIU.
DR InterPro; IPR001364; Hemagglutn.
DR Pfam; PF00509; Hemagglutinin; 1.
DR ProDom; PD000225; Hemagglutn; 1.
DR Envelope protein; Hemagglutinin; Glycoprotein.
FT NON_TER 1 1
FT CHAIN 1 328 HEMAGGLUTININ HA1 CHAIN.
FT CHAIN 330 550 HEMAGGLUTININ HA2 CHAIN.
FT CARBOHYD 8 8 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 285 285 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 137 137 K -> N (IN PIR DATA BANK).
SQ SEQUENCE 550 AA; 61659 MW; A107023ACC9CC353 CRC64;

Qy 6 QNTLKLATGKGGPKYKQNTLKLATGKGGV 35
|| | : | ||||| ||||| : |
Db 295 QNVNKITYG-ACPKYKQNTLKLATGMNV 323

Query Match 40.1%; Score 81.5; DB 1; Length 550;
Best Local Similarity 63.3%; Pred. No. 0.0062;
Matches 19; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

```

```

RL Virology 159:109-119(1987).
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO CELL RECEPTORS AND FOR INITIATING INFECTION.
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M16738; AAA43144.1; -.
DR PIR; B27813; HMIV80.
DR HSSP; P03437; 2VIU.
DR InterPro; IPR001364; Hemagglutn.
DR Pfam; PF00509; Hemagglutinin; 1.
DR ProDom; PD000225; Hemagglutn; 1.
DR Envelope protein; Hemagglutinin; Glycoprotein.
FT NON_TER 1 1
FT CHAIN 1 328 HEMAGGLUTININ HA1 CHAIN.
FT CHAIN 330 550 HEMAGGLUTININ HA2 CHAIN.
FT CARBOHYD 8 8 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 285 285 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 137 137 K -> N (IN PIR DATA BANK).
SQ SEQUENCE 550 AA; 61659 MW; A107023ACC9CC353 CRC64;

Query Match 40.1%; Score 81.5; DB 1; Length 550;
Best Local Similarity 63.3%; Pred. No. 0.0062;
Matches 19; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

Qy 6 QNTLKLATGKGGPKYKQNTLKLATGKGGV 35
|| | : | ||||| ||||| : |
Db 295 QNVNKITYG-ACPKYKQNTLKLATGMNV 323

RESULT 15
HEMA_IADH3
ID HEMA_IADH3 STANDARD; PRT; 550 AA.
AC P12584; Q84012; Q89793;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hemagglutinin precursor [Contains: Hemagglutinin HA1 chain; Hemagglutinin HA2 chain] (Fragment).
GN HA.
OS Influenza A virus (strain A/Duck/Hokkaido/33/80).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
CC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=11359;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87265458; PubMed=2440178;
RA Kida H., Kawaoka Y., Naeve C.W., Webster R.G.;
RT "Antigenic and genetic conservation of H3 influenza virus in wild ducks.";
RL Virology 159:109-119(1987).
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO CELL RECEPTORS AND FOR INITIATING INFECTION.
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS (HA1 AND HA2) LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M16738; AAA43144.1; -.
DR PIR; B27813; HMIV80.
DR HSSP; P03437; 2VIU.
DR InterPro; IPR001364; Hemagglutn.
DR Pfam; PF00509; Hemagglutinin; 1.
DR ProDom; PD000225; Hemagglutn; 1.
DR Envelope protein; Hemagglutinin; Glycoprotein.
FT NON_TER 1 1
FT CHAIN 1 328 HEMAGGLUTININ HA1 CHAIN.
FT CHAIN 330 550 HEMAGGLUTININ HA2 CHAIN.
FT CARBOHYD 8 8 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 285 285 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 137 137 K -> N (IN PIR DATA BANK).
SQ SEQUENCE 550 AA; 61659 MW; A107023ACC9CC353 CRC64;

Qy 6 QNTLKLATGKGGPKYKQNTLKLATGKGGV 35
|| | : | ||||| ||||| : |
Db 295 QNVNKITYG-ACPKYKQNTLKLATGMNV 323

Query Match 40.1%; Score 81.5; DB 1; Length 550;
Best Local Similarity 63.3%; Pred. No. 0.0062;
Matches 19; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

```

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; M16739; AAA43145.1; -  
DR PIR; C27813; HMIV33.  
DR HSSP; P03437; 2VIU.  
DR InterPro; IPR001364; Hemagglutn.  
DR Pfam; PF00509; Hemagglutinin; 1.  
DR ProDom; PD000225; Hemagglutn; 1.  
KW Envelope protein; Hemagglutinin; Glycoprotein.  
FT NON\_TER 1 1  
FT CHAIN 1 328 HEMAGGLUTININ HA1 CHAIN.  
FT CHAIN 330 550 HEMAGGLUTININ HA2 CHAIN.  
FT CARBOHYD 8 8 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 285 285 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 550 AA; 61577 MW; 6C30BF67CFDCB7DE CRC64;

Query Match 40.1%; Score 81.5; DB 1; Length 550;  
Best Local Similarity 63.3%; Pred. No. 0.0062;  
Matches 19; Conservative 2; Mismatches 8; Indels 1; Gaps 1;  
QY 6 QNTLKLATGKKGPKYVKQNTLKLATGKKG 35  
Db || | : | ||||| :  
295 QNVNKITYG-ACPKYVKQNTLKLATGMNRV 323

Search completed: December 4, 2002, 12:53:30  
Job time : 5.3333 secs

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: December 4, 2002, 12:49:06 ; Search time 15.5 seconds  
(without alignments)  
531.734 Million cell updates/sec

Title: US-09-362-731A-4  
Perfect score: 203  
Sequence: 1 PKYVKQNTLKLATGKPKYVKQNTLKLATGKKGVIIGIK 40

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : SPTREMBL.21.\*
- 1: sp\_archaea.\*
  - 2: sp\_bacteria.\*
  - 3: sp\_fungi.\*
  - 4: sp\_human.\*
  - 5: sp\_invertebrate.\*
  - 6: sp\_mammal.\*
  - 7: sp\_mbc.\*
  - 8: sp\_organelle.\*
  - 9: sp\_phage.\*
  - 10: sp\_plant.\*
  - 11: sp\_rodent.\*
  - 12: sp\_virus.\*
  - 13: sp\_vertibrate.\*
  - 14: sp\_unclassified.\*
  - 15: sp\_rvirus.\*
  - 16: sp\_bacteriap.\*
  - 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	85.5	42.1	344	12 Q86899	Q86899 equine infl
2	85.5	42.1	359	12 Q9YJW0	Q9YJW0 influenzavi
3	85.5	42.1	565	12 Q82559	Q82559 influenza a
4	83.5	41.1	344	12 Q82847	Q82847 influenza a
5	82.5	40.6	332	12 Q56964	Q56964 influenza a
6	82.5	40.6	335	12 Q91205	Q91205 influenza a
7	82.5	40.6	342	12 Q56963	Q56963 influenza a
8	82.5	40.6	343	12 Q9YJV7	Q9YJV7 influenzavi
9	82.5	40.6	344	12 Q9Q6P1	Q9Q6P1 equine infl
10	82.5	40.6	344	12 Q9Q6P0	Q9Q6P0 equine infl
11	82.5	40.6	344	12 Q9Q6N7	Q9Q6N7 equine infl
12	82.5	40.6	344	12 Q9Q6N6	Q9Q6N6 equine infl
13	82.5	40.6	344	12 Q9Q6N5	Q9Q6N5 equine infl
14	82.5	40.6	344	12 Q82843	Q82843 influenza a
15	82.5	40.6	344	12 Q82844	Q82844 influenza a
16	82.5	40.6	344	12 Q82846	Q82846 influenza a

17	82.5	40.6	344	12 Q82848	Q82848 influenza a
18	82.5	40.6	346	12 Q56857	Q56857 equine infl
19	82.5	40.6	346	12 Q56858	Q56858 equine infl
20	82.5	40.6	353	12 O12294	O12294 equine infl
21	82.5	40.6	357	12 Q56856	Q56856 equine infl
22	82.5	40.6	360	12 Q9YJV8	Q9YJV8 influenzavi
23	82.5	40.6	360	12 Q9YJV3	Q9YJV3 influenzavi
24	82.5	40.6	362	12 Q9YJV9	Q9YJV9 influenzavi
25	82.5	40.6	362	12 Q9YJV6	Q9YJV6 influenzavi
26	82.5	40.6	363	12 Q9YJV4	Q9YJV4 influenzavi
27	82.5	40.6	365	12 Q56961	Q56961 influenza a
28	82.5	40.6	365	12 Q56962	Q56962 influenza a
29	82.5	40.6	565	12 Q82792	Q82792 influenzavi
30	82.5	40.6	565	12 Q82793	Q82793 influenzavi
31	82.5	40.6	585	12 Q67103	Q67103 influenza a
32	82.5	40.6	585	12 Q67104	Q67104 influenza a
33	82.5	40.6	565	12 Q67105	Q67105 influenza a
34	82.5	40.6	565	12 Q67106	Q67106 influenza a
35	82.5	40.6	565	12 Q67107	Q67107 influenza a
36	82.5	40.6	565	12 Q86639	Q86639 equine infl
37	82.5	40.6	585	12 Q9WA90	Q9WA90 influenzavi
38	81.5	40.1	328	12 Q90384	Q90384 influenza a
39	81.5	40.1	329	12 Q9YS47	Q9YS47 influenza a
40	81.5	40.1	329	12 Q9YS46	Q9YS46 influenza a
41	81.5	40.1	329	12 Q67024	Q67024 influenzavi
42	81.5	40.1	329	12 Q67025	Q67025 influenzavi
43	81.5	40.1	329	12 Q67346	Q67346 influenzavi
44	81.5	40.1	329	12 Q67351	Q67351 influenzavi
45	81.5	40.1	329	12 Q67352	Q67352 influenza a

ALIGNMENTS

RESULT 1

Q86899 ID Q86899 PRELIMINARY; PRT; 344 AA.  
AC Q86899; (TREMREL. 01, Created)  
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)  
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)  
DE Hemagglutinin (Fragment).  
GN HAL.  
OS Equine influenza virus H3N8.  
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;  
OC Influenza A viruses; Equine influenza virus.  
OX NCBI\_TaxID=31660;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95159661; PubMed=7545975;  
RA Oxburgh L., Berg M., Klingeborn B., Emmoth E., Linne T.;  
RT "Evolution of H3N8 equine influenza virus from 1963 to 1991.";  
RL Virus Res. 34:153-165(1994).  
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO  
CC CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).  
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS  
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.  
DR EMBL; S77429; AAB33340.2; -.  
DR HSSP; P03437; IHGE.  
DR InterPro; IPR001364; Hemagglutn.  
DR Pfam; PF00509; Hemagglutinin; 1.  
DR ProDom; PD000225; Hemagglutin; 1.  
KW Envelope protein; Glycoprotein; Hemagglutinin.  
FT NON\_TER 344 344  
SQ SEQUENCE 344 AA; 38284 MW; C885938341149DA6 CRC64;

Query Match 42.1%; Score 85.5; DB 12; Length 344;  
Best Local Similarity 63.3%; Pred. No. 0.0061;  
Matches 19; Conservative 3; Mismatches 7; Indels 1; Gaps 1;  
QY 6 QNTLKLATGKPKYVKQNTLKLATGKGV 35  
|| | : | | | : | | | | | | | | | | : |

```

DR EMBL; U58195; AAB02560.1; -.
DR HSSP; P03437; 1HTM.
DR InterPro; IPR001364; Hemagglutn.
DR Pfam; PF00509; Hemagglutinin; 1.
DR PRINTS; PR00329; HEMAGGLUTN12.
DR ProDom; PD000225; Hemagglutn; 1.
DR Envelope protein; Glycoprotein; Hemagglutinin.
SQ SEQUENCE 565 AA; 63633 MW; E8412D13945FA424 CRC64;

Query Match 42.1%; Score 85.5; DB 12; Length 565;
Best Local Similarity 63.3%; Pred. No. 0.011;
Matches 19; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

QY 6 QNTLKLATGKGGPKYVKONTLKLATGKKG 35
   || | : | | | : | | | | | | | | : |
Db 310 QNVNKITYG-KCPKYIRQNTLKLATGMRNV 338

RESULT 4
Q82847 PRELIMINARY; PRT; 344 AA.
ID AC Q82847
AC Q82847;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE HAY subunit of haemagglutinin (Fragment).
GN HAI.
OS Influenza A virus (A/eq/Newmarket/93/(H3N8)).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=159470;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=96203953; PubMed=8627254;
RX Dally J.M., Lai A.C., Binns M.M., Chambers T.M., Barrandeguy M.,
RA Mumford J.A.;
RT "Antigenic and genetic evolution of equine H3N8 influenza A viruses.";
RL J. Gen. Virol. 77:661-671(1996).
RN [2]
RN SEQUENCE FROM N.A.
RP Dally J.M.;
RA Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RP Yates P.;
RA Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC -1- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR EMBL; X85089; CAA59416.3; -.
DR HSSP; P03437; 2V10.
DR InterPro; IPR001364; Hemagglutn.
DR Pfam; PF00509; Hemagglutinin; 1.
DR PRINTS; PR00329; HEMAGGLUTN12.
DR ProDom; PD000225; Hemagglutn; 1.
DR Envelope protein; Glycoprotein; Hemagglutinin.
KW NON_TER 344
FT SEQUENCE 344 AA; 38270 MW; 2EA462241BD40F16 CRC64;

Query Match 41.1%; Score 83.5; DB 12; Length 344;
Best Local Similarity 60.0%; Pred. No. 0.011;
Matches 18; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 6 QNTLKLATGKGGPKYVKONTLKLATGKKG 35
   || | : | | | : | | | | | | | | : |
Db 310 QNVNKITYG-KCPKYIRQNTLKLATGMRNV 338

RESULT 5
O56964 PRELIMINARY; PRT; 332 AA.
ID O56964
DE O56964

```





DR	ProDom; PD000225; Hemagglutn; 1.
KW	Envelope protein; Glycoprotein; Hemagglutinin; Signal.
FT	SIGNAL 1 15 POTENTIAL
FT	CHAIN 16 >344 HEMAGGLUTININ.
FT	NON_TER 344 344
SQ	SEQUENCE 344 AA; 38270 MW; 2A3B7D226DA5880D CRC64;
Query Match 40.6%; Score 82.5; DB 12; Length 344;	
Best Local Similarity 60.0%; Pred.No. 0.015; Indels 1; Gaps 1;	
Matches 18; Conservative 4; Mismatches 7;	
Qy	6 QNTLKLTATGKKPKYVRONTLKLATGKKGV 35     :       :       :
Dd	310 QNVNKVTYG-KCPKYIRQNTLKLATGMNRV 338
RESULT 10	
Q9Q6P0	PRELIMINARY; PRT; 344 AA.
ID	Q9Q6P0
AC	Q9Q6P0;
DT	01-MAY-2000 (Tremblrel. 13, Created)
DT	01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT	01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE	Hemagglutinin precursor (Fragment).
GN	HA1.
OS	Equine influenza virus.
OC	Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC	Influenza A viruses.
OX	NCBI_TaxID=11312;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=A/EQ/SASKATOON/L/90;
RX	MEDLINE=21395189; PubMed=11504416;
RA	Lai A.C.K., Chambers T.M., Holland R.E. Jr., Morley P.S., Haines D.M.,
RA	Townsend H.G., Barrandeguy M.;
RT	"Diverged evolution of recent equine-2 influenza (H3N8) viruses in the
RT	Western Hemisphere.";
RL	Arch. Virol. 146:1063-1074(2001).
CC	-I- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC	CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC	-I- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC	(HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC	-I- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
CC	EMBL: AF197243; AAF22347.1; .
DR	HSP: P03437; 2VIU.
DR	InterPro: IPR001364; Hemagglutn.
DR	Pfam: PF00509; Hemagglutinin; 1.
DR	PRINTS: PR00329; HEMAGGLUTN12.
DR	ProDom; PD000225; Hemagglutin; 1.
KW	Envelope protein; Glycoprotein; Hemagglutinin; Signal.
FT	SIGNAL 1 15 POTENTIAL.
FT	CHAIN 16 >344 HEMAGGLUTININ.
FT	NON_TER 344 344
SQ	SEQUENCE 344 AA; 38322 MW; AD11B2D4EDA87F0A CRC64;
Query Match 40.6%; Score 82.5; DB 12; Length 344;	
Best Local Similarity 60.0%; Pred.No. 0.015; Indels 1; Gaps 1;	
Matches 18; Conservative 4; Mismatches 7;	
Qy	6 QNTLKLTATGKKPKYVRONTLKLATGKKGV 35     :       :       :
Dd	310 QNVNKVTYG-KCPKYIRQNTLKLATGMNRV 338
RESULT 11	
Q9Q6N7	PRELIMINARY; PRT; 344 AA.
ID	Q9Q6N7
AC	Q9Q6N7;
DT	01-MAY-2000 (Tremblrel. 13, Created)
DT	01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT	01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE	Hemagglutinin precursor (Fragment).
GN	HA1.



```
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=159467;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A/EQ/ARUNDEL/91/(H3N8);
RX MEDLINE=96203953; PubMed=8627254;
RA Daly J.M., Lai A.C., Binns M.M., Chambers T.M., Barrandeguy M.,
RA Mumford J.A.;
RT "Antigenic and genetic evolution of equine H3N8 influenza A viruses.";
RL J. Gen. Virol. 77:661-671(1996).
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR EMBL; X85085; CAA59413.2; -
DR HSSP; P03437; 2VIU.
DR InterPro: IPR001364; Hemagglutn.
DR Pfam: PF00509; Hemagglutinin; 1.
DR PRINTS; PR00329; HEMAGGLUTN12.
DR ProDom; PD000225; Hemagglutn; 1.
DR Envelope protein; Glycoprotein; Hemagglutinin.
KW NON_TER 344 344
FT SEQUENCE 344 AA; 38321 MW; B6BDCF71A7A3138E CRC64;
SQ

Query Match 40.6%; Score 82.5; DB 12; Length 344;
Best Local Similarity 60.0%; Pred. NO. 0.015;
Matches 18; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 6 ONTLKLATGKKGPKYVKONTLKLATGKKG 35
  || | : | | | : | | | | | | | | | |
DB 310 QNVNKVITYG-KCPKYIRONTLKLATGMRNV 338

RESULT 15
Q82844 Q82844 PRELIMINARY; PRT; 344 AA.
AC Q82844;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE HAY subunit of haemagglutinin (Fragment).
GN HAI.
OS Influenza A virus (A/eq/Ella/89/(H3N8)).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=159468;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A/EQ/ELLA/89/(H3N8);
RX MEDLINE=96203953; PubMed=8627254;
RA Daly J.M., Lai A.C., Binns M.M., Chambers T.M., Barrandeguy M.,
RA Mumford J.A.;
RT "Antigenic and genetic evolution of equine H3N8 influenza A viruses.";
RL J. Gen. Virol. 77:661-671(1996).
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR EMBL; X85086; CAA59413.2; -
DR HSSP; P03437; 2VIU.
DR InterPro: IPR001364; Hemagglutn.
DR Pfam; PF00509; Hemagglutinin; 1.
DR PRINTS; PR00329; HEMAGGLUTN12.
DR ProDom; PD000225; Hemagglutn; 1.
DR Envelope protein; Glycoprotein; Hemagglutinin.
KW NON_TER 344 344
FT SEQUENCE 344 AA; 38255 MW; 4C7C3192826633F9 CRC64;
SQ

Query Match 40.6%; Score 82.5; DB 12; Length 344;
Best Local Similarity 60.0%; Pred. NO. 0.015;
Matches 18; Conservative 4; Mismatches 7; Indels 1; Gaps 1;
```

```
QY 6 ONTLKLATGKKGPKYVKONTLKLATGKKG 35
  || | : | | | : | | | | | | | | | |
DB 310 QNVNKVITYG-KCPKYIRONTLKLATGMRNV 338

Search completed: December 4, 2002, 12:52:59
Job time : 16.5 secs
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 12:49:05 : Search time 19.8333 Seconds  
(without alignments)  
268.741 Million cell updates/sec

Title: US-09-362-731A-4  
Perfect score: 203  
Sequence: 1 PKYVKQNTLKLATGKPKYVKQNTLKLATGKGVIGIK 40

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : A\_Genesec.101002.\*  
1: /SIDS2/gcgdata/genesec/genesecp-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/genesec/genesecp-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/genesec/genesecp-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/genesec/genesecp-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/genesec/genesecp-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/genesec/genesecp-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/genesec/genesecp-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/genesec/genesecp-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/genesec/genesecp-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/genesec/genesecp-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/genesec/genesecp-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/genesec/genesecp-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/genesec/genesecp-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/genesec/genesecp-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/genesec/genesecp-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/genesec/genesecp-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/genesec/genesecp-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/genesec/genesecp-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/genesec/genesecp-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/genesec/genesecp-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/genesec/genesecp-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/genesec/genesecp-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/genesec/genesecp-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	203	100.0	40	21 AAY82635	Influenza A virus
2	85.5	42.1	565	8 AAF07011	Equine influenza v
3	85.5	42.1	565	11 AAR04943	Equine hemagglutin
4	85.5	42.1	565	19 AAW44946	EIV Fontainebleau
5	82.5	40.6	565	21 AAY70056	Wild type equine i
6	82.5	40.6	565	21 AAY70057	Cold-adapted equin
7	81.5	40.1	347	15 AAR63591	Stem region of A2/
8	81.5	40.1	386	5 AAP40615	Sequence of the X-
9	81.5	40.1	566	15 AAR63590	Full length H3N2 i
10	81.5	40.1	566	23 ABB05767	Influenza A/Udorn/

11	81.5	40.1	566	23 ABB05774	Influenza A/Udorn/
12	80.5	39.7	19	21 AAY99172	HLA class II bindi
13	80.5	39.7	566	19 AAW68406	SIV strain H3N2 ha
14	79.5	39.2	72	22 AAB46190	Tetanus toxoid epi
15	79.5	39.2	136	22 AAB49089	Anyloid beta tetan
16	77	37.9	19	21 AAB13854	Influenza virus ha
17	77	37.9	23	15 AAR60857	Influenza virus T-
18	77	37.9	23	15 AAR60860	Polyoxime COSM com
19	77	37.9	23	19 AAW59273	Haemagglutinin hea
20	77	37.9	24	19 AAW59268	Haemagglutinin hea
21	77	37.9	25	7 AAP60885	Synthetic peptide
22	77	37.9	25	19 AAW68315	MHC binding peptid
23	77	37.9	25	19 AAW63054	Influenza A haemag
24	77	37.9	25	20 AAY29714	Influenza virus an
25	77	37.9	25	21 AAY89679	Core polypeptide f
26	77	37.9	25	21 AAY68192	Altered MHC determ
27	77	37.9	25	21 AAY52846	Altered MHC determ
28	77	37.9	25	22 ABB01087	Viral DPl78/107-li
29	77	37.9	25	22 ABB02536	Viral core polypep
30	77	37.9	25	22 AAU13633	DPl78-like/DPl07-1
31	77	37.9	25	22 AAB78080	Core polypeptide T
32	77	37.9	25	22 AAB58607	Altered MHC determ
33	77	37.9	26	7 AAP60887	Synthetic peptide
34	75.5	37.2	570	18 AAW01669	Influenza A/Beilin
35	75.5	37.2	570	20 AAW75441	Influenza virus A/
36	75.5	37.2	571	22 AAE04951	Influenza virus A/
37	75.5	37.2	571	18 AAW01673	Influenza A/Shando
38	75.5	37.2	571	18 AAW01676	Influenza A/Johann
39	75.5	37.2	571	20 AAW75445	Influenza virus A/
40	75.5	37.2	571	20 AAW75448	Influenza virus A/
41	75.5	37.2	571	22 AAE04955	Influenza virus A/
42	75.5	37.2	571	22 AAE04958	Influenza virus A/
43	74	36.5	347	23 AAW76670	Influenza A virus
44	73	36.0	16	21 AAY99053	HLA class II bindi
45	73	36.0	18	23 AAE22829	Influenza virus ha

#### ALIGNMENTS

RESULT 1  
AAY82635  
ID AAY82635 standard; peptide: 40 AA.  
XX  
AC AAY82635;  
XX  
DT 07-AUG-2000 (first entry)  
XX  
DE Influenza A virus T cell epitope and Der p1 B cell epitope peptide.  
XX  
KW T cell epitope; B cell epitope; allergen; antigenic;  
KW antiallergic; antiasthmatic; antinflammatory; dermatological;  
KW immunosuppressive; vaccine; rhinitis; sinusitis; bronchial asthma;  
KW atopic dermatitis; acute urticaria; chronic urticaria;  
KW gastro-intestinal syndrome; food allergen; oro-pharyngeal syndrome;  
KW anaphylactic reaction; drug hypersensitivity; allergic reaction.  
XX  
OS Dermatophagoides pteronyssinus.  
OS Influenza virus.  
OS Synthetic.  
XX  
PN WO200006694-A2.  
XX  
PD 10-FEB-2000.  
XX  
PF 20-JUL-1999; 99WO-BE00092.  
XX  
PR 30-JUL-1998; 98EP-0870167.  
XX  
PA (UNIO ) UCB SA.  
XX  
PI Saint-Remy J, Jacquemin M;  
XX

DR WPI; 2000-422470/36.  
 XX New compound for prevention and treatment of allergies comprises at  
 PT least one allergen antigenic determinant recognized by a B cell and at  
 PT least one antigenic determinant which does not trigger T cell  
 PT activation.  
 XX  
 PS Claim 8; Page 35; 50pp; English.  
 XX  
 CC The present invention describes a compound (I) for the prevention and/or  
 CC treatment of allergy. The compound comprises at least one allergen  
 CC antigenic determinant (i) recognised by a B cell or an antibody secreted  
 CC by a B cell of a non-atopic individual and at least one antigenic  
 CC determinant (ii) different from the allergen that triggers T cell  
 CC activation. (i) has antiallergic, antialsthmatic, antiinflammatory,  
 CC dermatological and immunosuppressive activities, and can be used in a  
 CC treat and/or prevent allergies or a disease of allergic origin,  
 CC especially hypersensitivities. These include rhinitis, sinusitis,  
 CC bronchial asthma, atopic dermatitis, some forms of acute and chronic  
 CC urticaria, gastro-intestinal syndromes associated with the ingestion of  
 CC food allergens, oro-pharyngeal syndrome, anaphylactic reactions  
 CC associated with drug hypersensitivities and/or a mixture of these. The  
 CC use of (I) in the treatment of allergic conditions avoids the need for  
 CC drug treatment, which often causes undesirable side-effects. Also, prior  
 CC art drug therapies alleviate symptoms, but do not influence their  
 CC causes, however (i) actually combats the cause of an allergic reaction.  
 CC The present sequence represents a specifically claimed compound peptide  
 CC sequence from the present invention.  
 XX  
 SQ Sequence 40 AA;  
 Query Match 100.0%; Score 203; DB 21; Length 40;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-20;  
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 PKYVKNTLKLATGKPKYVKYKNTLKLATGKKGVIIGIK 40  
 DB 1 PKYVKNTLKLATGKPKYVKYKNTLKLATGKKGVIIGIK 40  
 RESULT 2  
 AAP70711  
 ID AAP70711 standard; protein; 565 AA.  
 AC  
 XX  
 AC AAP70711;  
 XX  
 DT 08-MAR-1991 (first entry)  
 XX  
 DE Equine influenza virus strain H3N8 (EIV-A2) haemagglutinin H3 gene  
 DE product.  
 DE  
 KW HA; vaccine; Vaccina.  
 XX  
 OS Equine influenza virus.  
 XX  
 FH Key Location/Qualifiers  
 FT Protein 17..343  
 FT /label= HA 1 protein  
 FT Protein 345..565  
 FT /label= HA 2 protein  
 XX  
 PN W08607593-A.  
 XX  
 PD 31-DEC-1986.  
 XX  
 XX 20-JUN-1986; 86WO-US01343.  
 PF  
 XX 20-JUN-1985; 85US-0747020.  
 PR  
 XX (BIOT-) BIOTECHN RES PARTNER.  
 PA  
 XX Dale B, Cordell B;  
 PI

XX WPI; 1987-007191/01.  
 DR N-PSDB; AAN71067.  
 XX  
 PT Preventing equine influenza virus infection - using recombinant  
 PT vaccines produced using DNA sequences encoding haemagglutinin and  
 PT neuraminidase glyco:proteins  
 XX  
 PS Disclosure; Fig 2; 63pp; English.  
 XX  
 CC Peptides derived from the haemagglutinin H7 and H3 and neuraminidase  
 CC N7 and N8 genes may be used to derive antigenic peptides useful in  
 CC vaccination against equine influenza virus infection.  
 CC Abs raised to the peptides may be used in diagnosis of the infection  
 CC and construction of probes to mutated forms of the virus.  
 XX  
 SQ Sequence 565 AA;  
 Query Match 42.1%; Score 85.5; DB 8; Length 565;  
 Best Local Similarity 63.3%; Pred. No. 0.0029;  
 Matches 19; Conservative 3; Mismatches 7; Indels 1; Gaps 1;  
 QY 6 QNTLKLATGKPKYVKYKNTLKLATGKKGCV 35  
 DB 310 QNVNKVYTG-KCPYIKNTLKLATGMRNV 338  
 RESULT 3  
 AAR04943  
 ID AAR04943 standard; protein; 565 AA.  
 AC  
 XX  
 AC AAR04943;  
 XX  
 DT 02-OCT-1990 (first entry)  
 XX  
 DE Equine hemagglutinin H3 (EIV-A2).  
 XX  
 KW Recombinant vaccines; equine influenza virus; haemagglutinin; H3;  
 KW neuraminidase; N8.  
 XX  
 OS Equine influenza virus.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 102..1182  
 FT /label=N-terminal HA1 50kD portion  
 FT Region 1186..1748  
 FT /label=C-terminal HA2 27kD portion  
 XX  
 PN US4920213-A.  
 XX  
 PD 24-APR-1990.  
 XX  
 XX 21-JUL-1986; 86US-0888250.  
 PF  
 XX 20-JUN-1985; 85US-0747020.  
 PR  
 XX 21-JUL-1986; 86US-0888250.  
 PR  
 XX (BIOT-) BIOTECH RES PARTNER.  
 PA  
 XX Dale B, Cordell B;  
 PI  
 DR WPI; 1990-163647/21.  
 DR P-PSDB; AAR04943.  
 XX  
 XX Recombinant vaccines against equine influenza virus - produced using DNA  
 XX sequences encoding haemagglutinin and neuraminidase glycoprotein(s).  
 PS Disclosure; ; 27pp; English.  
 XX  
 CC There are nine neuraminidase (NA) subtypes and twelve haemagglutinin  
 CC (HA) subtypes. The strain carrying H3N8 glycoproteins is designated  
 CC equine influenza virus (EVI)-A2. The cDNA sequences will be useful in  
 CC the construction of diagnostic probes for the disease and of probes for



```

DT 05-JUN-2000 (first entry)
XX Cold-adapted equine influenza virus H3N8 haemagglutinin protein.
DE Haemagglutinin protein; modified HA protein; horse; cold-adaptation;
KW reassortant virus; temperature sensitivity; dominant interference;
KW attenuation; antiviral; vaccine; prevention; treatment;
KW influenza A virus infection.
XX
XX Equine influenza virus H3N8.
OS WO200009702-A1.
XX
XX 24-FEB-2000.
PD
XX 12-AUG-1999; 99WO-US18583.
XX
XX 13-AUG-1998; 98US-0133921.
XX
XX (UYPI-) UNIV PITTSBURGH.
XX
XX Dowling PW, Youngner JS;
PI
XX
XX WPI; 2000-224339/19.
DR N-PSDB; AAZ50976.
XX
XX New cold-adapted equine influenza viruses and reassortant viruses used
PT as vaccines for treating influenza infections in animals, particularly
PT horses, have a phenotype such as temperature sensitivity or dominant
PT interference.
XX
XX Claim 10; Page 93-95; 127pp; English.
XX
XX The patent discloses experimentally generated cold-adapted equine
CC influenza viruses, and reassortant influenza A viruses comprising at
CC least one genome segment of the cold-adapted virus, which confers at
CC least one identifying phenotype selected from cold-adaptation,
CC temperature sensitivity, dominant interference and attenuation.
CC These viruses are used in therapeutic compositions e.g. vaccines for
CC preventing or treating infections caused by influenza A viruses in
CC animals, particularly horses. The present sequence is a
CC cold-adapted equine influenza virus H3N8 haemagglutinin (HA)
CC protein denoted as PeicelHA565. This sequence is a modified form of
CC the wild type sequence.
XX
XX Sequence 565 AA;
SQ
Query Match 40.6%; Score 82.5; DB 21; Length 565;
Best Local Similarity 60.0%; Pred. No. 0.0073;
Matches 18; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 6 ONTLKLTATGKGGPKYVKONTLKLATGKGV 35
|| | : | | | : | | | | | | | : |
Db 310 QNVNKTGTYG-KCPKYIROTNTLKLATGMNV 338

RESULT 7
AAR63591
ID AAR63591 standard; Protein; 347 AA.
XX
AC AAR63591;
XX
XX 22-JUN-1995 (first entry)
DT
XX Stem region of A2/Aichi/2/68 influenza virus.
DE
XX Conserved peptide; stem region; haemagglutinin; HA; H1N1; H2N2;
KW subtype; human; influenza A virus; immunogenic artificial peptide;
KW antigen; vaccine; infection.
XX
XX Human influenza A virus.
OS
XX Key Location/Qualifiers
FH

QY 6 ONTLKLTATGKGGPKYVKONTLKLATGKGV 35
|| | : | | | : | | | | | | | : |
Db 310 QNVNKTGTYG-KCPKYIROTNTLKLATGMNV 338

RESULT 8
AAP40615
ID AAP40615 standard; protein; 386 AA.
XX
AC AAP40615;
XX
XX 12-FEB-1992 (first entry)
DT
XX
DE Sequence of the X-47 haemagglutinin (HA) molecule contg. the entire
DE HA1 and the amino-terminus of HA2.
XX
XX Vaccine; immunogen; antigen; diagnosis; therapy; influenza; virus.
KW
XX Influenza virus type A (H3N2).
OS
XX Key Location/Qualifiers
FH
XX Protein 17..346
FT /label= HA1
FT Protein 347..386
FT /label= HA2
FT Peptide 69..76
FT /note= "claimed"
FT Peptide 156..176

```

```

FT Peptide 1..15
FT /note= "Signal peptide"
FT Protein 16..347
FT /note= "Mature HA protein"
FT Peptide 114..118
FT /note= "Conserved peptide"
FT Peptide 172..183
FT /note= "Conserved peptide"
XX
XX EP621339-A.
PN
XX 26-OCT-1994.
XX
XX 20-APR-1994; 94EP-0302819.
XX
XX 20-APR-1993; 93JP-0115216.
XX 16-MAR-1994; 94JP-0070194.
XX
XX (TAKI ) TAKARA SHUZO CO LTD.
XX
XX Isegawa Y, Okuno Y, Sasao F, Ueda S;
PI
XX WPI; 1994-325949/41.
DR N-PSDB; AAQ72807.
XX
XX Human influenza-A virus haemagglutinin polypeptide(s) - useful in
PT influenza-A vaccine composition
PT
XX Claim 17; Page 60-61; 68pp; English.
XX
XX This sequence represents the stem region of the hemagglutinin (HA)
CC molecule of the A2/Aichi/2/68 strain of human influenza A virus. This
CC antigenic molecule contains the conserved peptides derived from the stem
CC region of the H1N1 and H2N2 subtypes of human influenza A virus.
CC Immunogenic polypeptides such as this are antigenically equivalent to
CC the stem region of the HA molecule of influenza A virus. This artificial
CC peptide may be used as a vaccine for prophylaxis of influenza A virus
CC infection.
XX
XX Sequence 347 AA;
SQ
Query Match 40.1%; Score 81.5; DB 15; Length 347;
Best Local Similarity 63.3%; Pred. No. 0.0057;
Matches 19; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

QY 6 ONTLKLTATGKGGPKYVKONTLKLATGKGV 35
|| | : | | | | | | | | | : |
Db 92 QNVNKTGTYG-ACPKYVKONTLKLATGMNV 120

RESULT 8
AAP40615
ID AAP40615 standard; protein; 386 AA.
XX
AC AAP40615;
XX
XX 12-FEB-1992 (first entry)
DT
XX
DE Sequence of the X-47 haemagglutinin (HA) molecule contg. the entire
DE HA1 and the amino-terminus of HA2.
XX
XX Vaccine; immunogen; antigen; diagnosis; therapy; influenza; virus.
KW
XX Influenza virus type A (H3N2).
OS
XX Key Location/Qualifiers
FH
XX Protein 17..346
FT /label= HA1
FT Protein 347..386
FT /label= HA2
FT Peptide 69..76
FT /note= "claimed"
FT Peptide 156..176

```





PD 03-JAN-2002.  
XX  
PF 21-JUN-2001; 2001WO-US19826.  
XX  
PR 23-JUN-2000; 2000US-213650P.  
XX  
PA (AMCY ) AMERICAN CYANAMID CO.  
XX  
XX Galarza JM, Latham TE;  
PI  
XX WPI; 2002-139923/18.  
DR  
DR N-PSDB; ABA93937.  
XX  
PT Polynucleotide encoding complete sequence of influenza A/Udorn/72 and  
PT polypeptide, useful in diagnosis and for generating new influenza A  
PT variant strains  
XX  
PS Claim 19; Page 64-65; 103pp; English.  
XX  
XX The present invention describes an isolated polynucleotide (I) having  
CC the complete sequence of the Influenza A/Udorn/72 (H3N2) strain in  
CC positive strand, antigenomic message sense. ABA93934 to ABA93944 encode  
CC the Influenza A/Udorn/72 (H3N2) strain proteins given in ABB05764 to  
CC ABB05774 from the present invention. (I) is useful for designing  
CC polymerase chain reaction (PCR) primers for use in a PCR assay to detect  
CC the presence of the corresponding virus segment in a sample or for  
CC designing and selecting peptides for use in an enzyme linked  
CC immunosorbant assay to detect the presence of the corresponding protein  
CC produced by that segment in a sample, hence is useful in diagnosis and  
CC may be modified by mutation to generate new influenza A variant strains.  
CC ABA94945 to ABA94039 represent Influenza A/Udorn/72 (H3N2) strain  
CC sequencing primers, which are used in an example from the present  
CC invention.  
XX  
SQ Sequence 566 AA;  
Query Match 40.1%; Score 81.5; DB 23; Length 566;  
Best Local Similarity 63.3%; Pred. No. 0.01;  
Matches 19; Conservative 2; Mismatches 8; Indels 1; Gaps 1;  
QY 6 QNTLKLATGKGGPKYVKQNTLKLATGKGGV 35  
DB 311 QNVNKITYG-ACPKYVKQNTLKLATGMNRV 339  
RESULT 11  
ABB05774  
ID ABB05774 standard; Protein; 566 AA.  
XX  
AC ABB05774;  
XX  
XX 07-MAY-2002 (first entry)  
DT  
XX Influenza A/Udorn/72 (H3N2) Strain HA protein SEQ ID NO:22.  
DE  
XX Influenza A/Udorn/72 (H3N2) strain; Influenzavirus A; diagnosis;  
KW Influenza A virus; genome.  
KW  
XX Influenzavirus A.  
OS  
XX WO200200884-A2.  
PN  
XX  
PD 03-JAN-2002.  
XX  
PF 21-JUN-2001; 2001WO-US19826.  
XX  
PR 23-JUN-2000; 2000US-213650P.  
XX  
PA (AMCY ) AMERICAN CYANAMID CO.  
XX  
XX Galarza JM, Latham TE;  
PI  
XX WPI; 2002-139923/18.  
DR

DR N-PSDB; ABA93944.  
XX  
PT Polynucleotide encoding complete sequence of influenza A/Udorn/72 and  
PT polypeptide, useful in diagnosis and for generating new influenza A  
PT variant strains  
XX  
PS Disclosure; Page 83-85; 103pp; English.  
XX  
XX The present invention describes an isolated polynucleotide (I) having  
CC the complete sequence of the Influenza A/Udorn/72 (H3N2) strain in  
CC positive strand, antigenomic message sense. ABA93934 to ABA93944 encode  
CC the Influenza A/Udorn/72 (H3N2) strain proteins given in ABB05764 to  
CC ABB05774 from the present invention. (I) is useful for designing  
CC polymerase chain reaction (PCR) primers for use in a PCR assay to detect  
CC the presence of the corresponding virus segment in a sample or for  
CC designing and selecting peptides for use in an enzyme linked  
CC immunosorbant assay to detect the presence of the corresponding protein  
CC produced by that segment in a sample, hence is useful in diagnosis and  
CC may be modified by mutation to generate new influenza A variant strains.  
CC ABA94945 to ABA94039 represent Influenza A/Udorn/72 (H3N2) strain  
CC sequencing primers, which are used in an example from the present  
CC invention.  
XX  
SQ Sequence 566 AA;  
Query Match 40.1%; Score 81.5; DB 23; Length 566;  
Best Local Similarity 63.3%; Pred. No. 0.01;  
Matches 19; Conservative 2; Mismatches 8; Indels 1; Gaps 1;  
QY 6 QNTLKLATGKGGPKYVKQNTLKLATGKGGV 35  
DB 311 QNVNKITYG-ACPKYVKQNTLKLATGMNRV 339  
RESULT 12  
AAY99172  
ID AAY99172 standard; Peptide; 19 AA.  
XX  
AC AAY99172;  
XX  
XX 07-AUG-2000 (first entry)  
DT  
XX HLA class II binding antigen epitope peptide #361.  
DE  
XX Human leucocyte antigen; HLA class II; antigen epitope; pharmaceutical;  
KW immune response; chronic viral disease; cancer; autoimmune disease;  
KW rheumatoid arthritis; multiple sclerosis; myasthenia gravis; AIDS;  
KW allograft rejection; allergy; Lyme disease; hepatitis; prostate cancer;  
KW glomerulonephritis; food hypersensitivity; malaria.  
XX  
OS Unidentified.  
XX  
XX WO9961916-A1.  
PN  
XX 02-DEC-1999.  
PD  
XX 28-MAY-1999; 99WO-US12066.  
PF  
XX 29-MAY-1998; 98US-0087192.  
PR  
XX (EPIM-) EPIMMUNE INC.  
PA  
XX Sette A, Southwood S, Sidney J;  
PI  
XX WPI; 2000-097143/08.  
DR  
XX New compositions containing immunogenic peptide epitopes for various  
PT HLA class II DR molecules useful for inducing helper T cell response  
XX  
XX Claim 1; Page 46; 60pp; English.  
PS  
XX The present invention relates to a new pharmaceutical composition  
CC comprising a unit dose form of a peptide, or analogue, comprising an

CC epitope selected from those represented by peptides AAY98812-Y99339  
 CC which are derived from various antigens for various human leucocyte  
 CC antigen class DR molecules, representative of the world wide population.  
 CC The peptide/analogue binds to an HLA class II molecule at an IC-50 of  
 CC less than or equal to 1,000 nM. The pharmaceutical can be used to induce  
 CC a helper T cell response. The pharmaceutical focuses the immune response  
 CC towards selected determinants and could therefore be used in cases of  
 CC chronic viral diseases and cancer. Examples of diseases that can be  
 CC treated using the peptide containing pharmaceutical include autoimmune  
 CC diseases (rheumatoid arthritis, multiple sclerosis, and myasthenia  
 CC gravis), allograft rejection, allergies, Lyme disease, hepatitis,  
 CC post-streptococcal endocarditis or glomerulonephritis and food  
 CC hypersensitivities. The peptide epitopes can be used to enhance immune  
 CC responses against other immunogens administered with the peptides.  
 CC Diseases which can be treated using immunogenic mixtures include  
 CC prostate cancer, hepatitis B, hepatitis C, AIDS, renal carcinoma,  
 CC cervical carcinoma, lymphoma, and condyloma acuminatum. The peptides may  
 CC also be used to make monoclonal antibodies useful as potential diagnostic  
 CC or therapeutic agents. The peptides may also be useful as diagnostic  
 CC reagents, for example, to determine the susceptibility of an individual  
 CC to a treatment regimen. Also, the peptides may be used to predict which  
 CC individuals will be at substantial risk of developing chronic infection.  
 CC The selection of appropriate T and B cell epitopes should allow the  
 CC development of epitope based vaccines particularly towards conserved  
 CC epitopes of pathogens which are characterized by high sequence  
 CC variability such as HIV, HCV and Malaria.  
 XX  
 SQ Sequence 19 AA:  
 Query Match 39.7%; Score 80.5; DB 21; Length 19;  
 Best Local Similarity 73.1%; Pred. No. 0.00025;  
 Matches 19; Conservative 0; Mismatches 0; Indels 7; Gaps 1;  
 QY 3 YVKQNTLKLATGKGGPKYKQNTLKL 28  
 DB 1 YVKQNTLKL-----YVKQNTLKL 19  
 RESULT 13  
 AAW68406  
 ID AAW68406 standard; Protein: 566 AA.  
 XX  
 AC AAW68406;  
 XX  
 DT 28-OCT-1998 (first entry)  
 XX  
 DE SIV strain H3N2 haemagglutinin.  
 XX  
 KW Multivalent vaccine; pig; pathogen; respiratory disease; SIV; PRRSV; HCV;  
 KW digestive disease; Aujeszky's disease virus; pseudorabies virus; vaccine;  
 KW swine herpesvirus 1; swine influenza virus; hog cholera virus; vector;  
 KW porcine respiratory and reproductive syndrome virus; haemagglutinin;  
 KW SIRS; swine infertility and respiratory syndrome virus;  
 KW Actinobacillus pleuropneumoniae.  
 XX  
 OS Swine influenza virus.  
 XX  
 PN FR2751224-A1.  
 XX  
 PD 23-JAN-1998.  
 XX  
 PF 19-JUL-1996; 96FR-00093338.  
 XX  
 PR 19-JUL-1996; 96FR-00093338.  
 XX  
 PA (INWR ) RHONE MERIEUX SA.  
 XX  
 DR WPI; 1998-112824/11.  
 DR N-PSDB; AAV49298.  
 XX  
 PT Multi-valent polynucleotide vaccines against porcine pathogens -  
 PT consist of at least 3 plasmids able to express protective antigens  
 PT from specified viruses

XX Example 12; Fig 10; 63pp; French.  
 XX  
 CC The invention relates to a multivalent vaccine for protecting pigs  
 CC against several pathogens, especially pathogens associated with  
 CC respiratory and digestive diseases. The pathogens are especially  
 CC selected from Aujeszky's disease virus, swine influenza virus (SIV),  
 CC porcine respiratory and reproductive syndrome virus (PRRSV), hog  
 CC cholera virus (HCV) and Actinobacillus pleuropneumoniae. The vaccines  
 CC are preferably composed of polynucleotide sequences encoding 3 antigens,  
 CC all as part of vectors. This sequence represents the SIV strain H3N2  
 CC haemagglutinin protein. The coding sequence was subcloned into the  
 CC plasmid pVR1012 to generate plasmid pPB144 for use in the vaccine.  
 XX  
 SQ Sequence 566 AA;  
 Query Match 39.7%; Score 80.5; DB 19; Length 566;  
 Best Local Similarity 60.0%; Pred. No. 0.014;  
 Matches 18; Conservative 3; Mismatches 8; Indels 1; Gaps 1;  
 QY 6 QNTLKLATGKGGPKYKQNTLKLATGKGGV 35  
 DB 311 QNVNKITYG-ACPKYKQNTLKLATGMRI 339  
 RESULT 14  
 AAB46190  
 ID AAB46190 standard; peptide: 72 AA.  
 XX  
 AC AAB46190;  
 XX  
 DT 04-APR-2001 (first entry)  
 XX  
 DE Tetanus toxoid epitope fusion construct #10.  
 XX  
 KW Amyloid deposit; APP; Abeta; brain; human; clearing response; neurotropic;  
 KW Fc receptor mediated phagocytosis; immunogenic response; neuroprotective;  
 KW amyloid precursor protein; Alzheimer's disease.  
 XX  
 OS Clostridium tetani.  
 XX  
 PN WO200072880-A2.  
 XX  
 PD 07-DEC-2000.  
 XX  
 PF 26-MAY-2000; 2000WO-US14810.  
 XX  
 PR 28-MAY-1999; 99US-0322289.  
 XX  
 PA (NEUR-) NEURALAB LTD.  
 XX  
 PI Schenk DB, Bard F, Vasquez NJ, Yednock T;  
 XX WPI; 2001-032104/04.  
 DR  
 XX Preventing or treating a disease associated with amyloid deposits,  
 PT especially Alzheimer's disease, comprises administering amyloid  
 PT specific antibody -  
 XX  
 PS Disclosure; Page 32; 143pp; English.  
 XX  
 CC This invention describes a novel method of preventing or treating a  
 CC disease associated with amyloid deposits of amyloid precursor protein  
 CC (APP) Abeta fragments in the brain of a patient, which comprises  
 CC administering to the patient: (a) an antibody that binds to Abeta, the  
 CC antibody binds to an amyloid deposit and induces a clearing response (Fc  
 CC receptor mediated phagocytosis) against it (b) a polypeptide containing  
 CC an N-terminal segment of at least residues 1-5 of Abeta; or (c) an agent  
 CC that induces an immunogenic response against residues 1-3 to 7-11 of  
 CC Abeta. The products of the invention have neurotropic and neuroprotective  
 CC activity. The method is also useful for monitoring a course of treatment  
 CC being administered to a patient e.g. active and passive immunization. The  
 CC methods are useful for prophylactic and therapeutic treatment of

CC Alzheimer's disease.  
XX  
SQ Sequence 72 AA; Length 72; DB 22; Mismatches 2; Indels 11; Gaps 1;  
Query Match 39.2%; Score 79.5; Pred. No. 0.0017;  
Best Local Similarity 52.8%;  
Matches 19; Conservative 2;  
QY 1 PRYVKQNTLKLATGKKGPK-----YVKQNT 25  
DB 8 PRYVKQNTLKLATEKKIAKMEKASSVFNQYIKANS 43  
RESULT 15  
AAB49089  
ID AAB49089 standard; Protein; 136 AA.  
XX  
AC AAB49089;  
XX  
DT 27-MAR-2001 (first entry)  
XX  
DE Amyloid beta tetanus toxoid/HA/CS fusion protein, SEQ ID NO:25.  
XX  
KW Amyloid disease; amyloid fibril deposition; amyloid plaque;  
KW immunogenic; antibody; vaccine; Alzheimer's disease;  
KW type 2 diabetes; reactive system amyloidosis;  
KW systemic senile amyloidosis; familial amyloidosis;  
KW transmissible spongiform encephalopathy; Creutzfeld-Jakob disease; Kuru;  
KW haemodialysis-associated beta-2-microglobulin deposition;  
KW amyloid beta peptide; universal T-cell epitope; neuroprotective.  
XX  
OS Chimeric - Homo sapiens.  
OS Chimeric - Clostridium tetani.  
OS Chimeric - Influenza virus.  
OS Chimeric - Plasmodium falciparum.  
XX  
FN WO200072876-A2.  
XX  
PD 07-DEC-2000.  
XX  
PF 01-JUN-2000; 2000WO-US15239.  
XX  
PR 01-JUN-1999; 99US-0137010.  
XX  
PA (NEUR-) NEURALAB LTD.  
XX  
PI Schenk DB;  
XX  
DR WPI; 2001-070921/08.  
XX  
XX Pharmaceutical composition comprising immunogen against amyloid  
PT component such as fibril peptide or protein, or antibody against  
PT amyloid component useful for treating amyloid diseases or amyloidoses -  
XX  
PS Disclosure; Page 46; 140pp; English.  
XX  
XX The invention relates to a novel pharmaceutical composition for  
CC preventing or treating a disease characterised by amyloid fibril  
CC deposits (amyloid plaques) in a patient. The pharmaceutical composition  
CC comprises an agent that will induce an immune response against an amyloid  
CC component, or an antibody or antibody fragment that binds to an amyloid  
CC component. The invention also relates to a method for determining  
CC the prognosis of a patient undergoing treatment for an amyloid disorder  
CC which involves measuring a patient serum amount of immunoreactivity  
CC against a selected amyloid component. A patient serum immunoreactivity  
CC of at least four times a base line serum immunoreactivity control level  
CC indicates a prognosis of improved status with respect to the disorder.  
CC The pharmaceutical compositions of the invention are useful for treating  
CC a wide variety of disorders characterised by amyloid fibril deposition in  
CC a patient. Such disorders include Alzheimer's disease characterised by  
CC amyloid beta peptide fibril deposits; type 2 diabetes characterised by  
CC islet amyloid protein peptide (IAPP, amylin) fibrils; reactive systemic  
CC amyloidosis associated with systemic inflammatory diseases (e.g.,  
CC rheumatoid arthritis, osteomyelitis, tuberculosis) characterised by AA  
CC fibrils derived from serum amyloid A protein (ApoSAA)); systemic senile  
CC amyloidosis and familial amyloid cardiomyopathy characterised by ATTR  
CC fibrils derived from transthyretin (TTR); transmissible spongiform  
CC encephalopathies (e.g. Creutzfeld-Jakob disease, Kuru) characterised by  
CC prion protein deposits; and beta-2-microglobulin deposits which form as  
CC a result of long term haemodialysis treatment. The present sequence  
CC represents an immunogenic fusion protein comprising an amyloid beta  
CC peptide fused to a universal T-cell epitope which may be used in a  
CC composition to treat or prevent Alzheimer's disease.  
XX  
SQ Sequence 136 AA;  
Query Match 39.2%; Score 79.5; DB 22; Length 136;  
Best Local Similarity 52.8%; Pred. No. 0.0035;  
Matches 19; Conservative 2; Mismatches 4; Indels 11; Gaps 1;  
QY 1 PRYVKQNTLKLATGKKGPK-----YVKQNT 25  
DB 8 PRYVKQNTLKLATEKKIAKMEKASSVFNQYIKANS 43  
Search completed: December 4, 2002, 12:51:16  
Job time : 22.0833 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 12:53:35 ; Search time 3.66667 Seconds  
(without alignments)  
177.189 Million cell updates/sec

Title: US-09-362-731a-4

Perfect score: 203

Sequence: 1 PKVVKQNTLKLATGCKPKYVVKNTLKLATGCKKGIIGIK 40

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA:\*

```
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	81.5	40.1	347	10	US-09-918-568-58
2	77	37.9	23	9	US-10-044-034-3
3	77	37.9	24	9	US-10-044-034-11
4	72	35.5	14	9	US-10-044-034-15
5	72	35.5	14	10	US-09-202-077-13
6	72	35.5	14	10	US-09-202-077-15
7	68	33.5	16	9	US-10-044-034-6
8	66.5	32.8	564	9	US-10-099-619-2
9	66	32.5	13	10	US-09-848-164-8
10	66	32.5	13	10	US-09-768-872-2
11	66	32.5	13	10	US-09-245-487B-23
12	66	32.5	13	10	US-09-756-983-8
13	66	32.5	13	10	US-09-756-983-14
14	63	31.0	17	9	US-10-044-034-5
15	62	30.5	13	9	US-09-987-137-7
16	61	30.0	12	9	US-10-044-034-4
17	52	25.6	348	10	US-09-918-568-50
18	50.5	24.9	1357	10	US-09-815-242-11997
19	50	24.6	10	9	US-10-044-034-27

```

20 50 24.6 347 10 US-09-415-277A-11 Sequence 11, Appl
21 49.5 24.4 1172 9 US-09-712-363-176 Sequence 176, Appl
22 48 23.6 1475 10 US-09-740-274-2 Sequence 2, Appl
23 47.5 23.4 207 12 US-10-043-142-11 Sequence 11, Appl
24 47.5 23.4 1083 9 US-10-108-605-77 Sequence 77, Appl
25 47 23.2 208 10 US-09-925-299-982 Sequence 982, Appl
26 47 23.2 330 10 US-09-815-242-10691 Sequence 10691, A
27 46 22.7 284 10 US-09-815-242-12973 Sequence 12973, A
28 46 22.7 412 10 US-09-741-669-331 Sequence 331, Appl
29 46 22.7 607 10 US-09-815-242-5825 Sequence 5825, A
30 46 22.7 1349 10 US-09-747-835A-52 Sequence 52, Appl
31 45.5 22.4 208 12 US-10-043-142-12 Sequence 12, Appl
32 45.5 22.4 291 10 US-09-925-301-1323 Sequence 1323, A
33 45.5 22.4 364 9 US-09-971-228-6 Sequence 6, Appl
34 45.5 22.4 364 10 US-09-842-316-3 Sequence 3, Appl
35 45.5 22.4 364 10 US-09-731-030A-15 Sequence 15, Appl
36 45.5 22.4 364 12 US-10-037-616-20 Sequence 20, Appl
37 45.5 22.4 749 9 US-10-062-730-2 Sequence 2, Appl
38 45.5 22.4 1165 9 US-10-075-460-6 Sequence 6, Appl
39 45.5 22.4 1165 10 US-09-887-052-2 Sequence 2, Appl
40 45.5 22.4 1165 10 US-09-887-052-4 Sequence 4, Appl
41 45.5 22.4 1165 10 US-09-887-052-6 Sequence 6, Appl
42 45 22.2 116 10 US-09-764-853-437 Sequence 437, Appl
43 45 22.2 137 10 US-09-864-761-45751 Sequence 45751, A
44 45 22.2 269 9 US-10-027-806-6 Sequence 6, Appl
45 45 22.2 348 10 US-09-864-761-37014 Sequence 37014, A

```

#### ALIGNMENTS

#### RESULT 1

US-09-918-568-58

; Sequence 58, Application US/09918568

; Patent No. US20020054882A1

; GENERAL INFORMATION:

; APPLICANT: Yoshinobu OKUNO et al.

; TITLE OF INVENTION: POLYPEPTIDES FOR USE IN GENERATING

; ANTI-HUMAN INFLUENZA VIRUS ANTIBODIES (AS AMENDED)

; NUMBER OF SEQUENCES: 58

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.

; STREET: 2033 K Street, N.W., #800

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20006

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: Wordperfect 5.1

; CURRENT APPLICATION NUMBER: US/09/918,568

; FILING DATE: 02-Aug-2001

; CLASSIFICATION: <Unknown>

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: 09/004,422

; FILING DATE: January 8, 1998

; APPLICATION NUMBER: 08/443,862

; FILING DATE: May 22, 1995

; APPLICATION NUMBER: 08/229,781

; FILING DATE: April 19, 1994

; APPLICATION NUMBER: 08/034,016

; FILING DATE: April 29, 1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Warren M. Cheek, Jr.

; REGISTRATION NUMBER: 33,367

; REFERENCE/DOCKET NUMBER: <Unknown>

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-721-8200

; TELEFAX: 202-721-8250

; TELEX: <Unknown>



```
; PRIOR FILING DATE: 1997-02-11
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptides
US-10-044-034-15

Query Match          35.5%; Score 72; DB 9; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKYVKQNTLKLATG 14
| | | | | | | | | |
Db 1 PKYVKQNTLKLATG 14

RESULT 5
US-09-202-077-13
; Sequence 13, Application US/09202077
; Patent No. US20020103335A1
; GENERAL INFORMATION:
; APPLICANT: OLDHAM, Keith
; APPLICANT: EDWARDS, Philip N.
; APPLICANT: LUKE, Richard W.
; APPLICANT: COTTON, Ronald
; TITLE OF INVENTION: PEPTIDE DERIVATIVES
; FILE REFERENCE: 1991-140
; CURRENT APPLICATION NUMBER: US/09/202,077
; PRIOR FILING DATE: 1999-05-26
; PRIOR APPLICATION NUMBER: GB 9611881.5
; PRIOR FILING DATE: 1996-06-07
; PRIOR APPLICATION NUMBER: GB 9622890.3
; PRIOR FILING DATE: 1996-11-02
; PRIOR APPLICATION NUMBER: GB 9701491
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: NP_BIND
; LOCATION: (1)
; OTHER INFORMATION: Biotin-Ahx-
; NAME/KEY: NP_BIND
; LOCATION: (14)
; OTHER INFORMATION: -OH
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: peptide
US-09-202-077-13

Query Match          35.5%; Score 72; DB 10; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKYVKQNTLKLATG 14
| | | | | | | | | |
Db 1 PKYVKQNTLKLATG 14

RESULT 6
US-09-202-077-15
; Sequence 15, Application US/09202077
; Patent No. US20020103335A1
; GENERAL INFORMATION:
; APPLICANT: OLDHAM, Keith
; APPLICANT: EDWARDS, Philip N.
```

```
; APPLICANT: LUKE, Richard W.
; APPLICANT: COTTON, Ronald
; TITLE OF INVENTION: PEPTIDE DERIVATIVES
; FILE REFERENCE: 1991-140
; CURRENT APPLICATION NUMBER: US/09/202,077
; PRIOR FILING DATE: 1999-05-26
; PRIOR APPLICATION NUMBER: GB 9611881.5
; PRIOR FILING DATE: 1996-06-07
; PRIOR APPLICATION NUMBER: GB 9622890.3
; PRIOR FILING DATE: 1996-11-02
; PRIOR APPLICATION NUMBER: GB 9701491
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: peptide
US-09-202-077-15

Query Match          35.5%; Score 72; DB 10; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKYVKQNTLKLATG 14
| | | | | | | | | |
Db 1 PKYVKQNTLKLATG 14

RESULT 7
US-10-044-034-6
; Sequence 6, Application US/10044034
; Patent No. US20020169264A1
; GENERAL INFORMATION:
; APPLICANT: JACKSON, DAVID C.
; APPLICANT: O'BRIEN-SIMPSON, NEIL M.
; APPLICANT: BROWN, LORENA E.
; APPLICANT: EDE, NICHOLAS J.
; APPLICANT: BRANDT, EVELYN R.
; APPLICANT: GOOD, MICHAEL F.
; TITLE OF INVENTION: POLYMERS INCORPORATING PEPTIDES
; FILE REFERENCE: FBRC:006
; CURRENT APPLICATION NUMBER: US/10/044,034
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: P05071
; PRIOR FILING DATE: 1997-02-11
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptides
US-10-044-034-6

Query Match          33.5%; Score 68; DB 9; Length 16;
Best Local Similarity 87.5%; Pred. No. 0.0011;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 14 GKKGPKYVKQNTLKL 29
| | | | | | | | | |
Db 1 GFLGPKYVKQNTLKL 16

RESULT 8
US-10-099-619-2
; Sequence 2, Application US/10099619
; Patent No. US20020168384A1
```

; GENERAL INFORMATION:
; APPLICANT: CLAESSENS, JOHANNES AJ
; APPLICANT: WALTER, FUCHS
; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS VACCINE
; FILE REFERENCE: 20010010US
; CURRENT APPLICATION NUMBER: US/10/099,619
; CURRENT FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: EP012009759
; PRIOR FILING DATE: 2001-03-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 564
; TYPE: PRT
; ORGANISM: AVIAN INFLUENZA VIRUS
US-10-099-619-2

Query Match 32.8%; Score 66.5; DB 9; Length 564;
Best Local Similarity 56.7%; Pred. No. 0.12;
Matches 17; Conservative 2; Mismatches 10; Indels 1; Gaps 1;

Qy 6 QNTLKLATGKKGPKYKQNTLKLATGKKG 35
Db 304 QNINRAVG-KCPRYVKQESLLLATGMKNV 332

RESULT 9
US-09-848-164-8
; Sequence 8, Application US/09848164
; Patent No. US20020034513A1
; GENERAL INFORMATION:
; APPLICANT: Rhode, Peter R.
; APPLICANT: Jiao, Jin-An
; APPLICANT: Burkhardt, Martin
; APPLICANT: Wong, Hing
; TITLE OF INVENTION: MHC COMPLEXES AND USES THEREOF
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dade International, Inc.
; STREET: 1717 Deerfield Road
; CITY: Deerfield
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60015
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/848,164
; FILING DATE: 03-May-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/067,615
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/382,454
; FILING DATE: 01-FEB-1995
; APPLICATION NUMBER: US 08/283,302
; FILING DATE: 29-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Pearson, Louise S.
; REGISTRATION NUMBER: 32,369
; REFERENCE/DOCKET NUMBER: STR-4665-CIP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (708) 267-5300
; TELEFAX: (708) 267-5376
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown

; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-848-164-8

Query Match 32.5%; Score 66; DB 10; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKYVKQNTLKLAT 13
Db 1 PKYVKQNTLKLAT 13

RESULT 10
US-09-768-872-2
; Sequence 2, Application US/09768872
; Patent No. US20020055466A1
; GENERAL INFORMATION:
; APPLICANT: Aharoni, Rina
; APPLICANT: Teitelbaum, Dvora
; APPLICANT: Arnon, Ruth
; APPLICANT: Sela, Michael
; APPLICANT: Fridkis-Harelli, Masha
; APPLICANT: Strominger, Jack
; TITLE OF INVENTION: Treatment of Autoimmune Conditions with Copolymer 1
; FILE REFERENCE: 1662/493762
; CURRENT APPLICATION NUMBER: US/09/768,872
; CURRENT FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: US 60/093,859
; PRIOR FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: US 60/101,825
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: US 60/102,960
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: US 60/106,350
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: US 60/108,184
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/123,675
; PRIOR FILING DATE: 1999-03-09
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: WordPerfect 8.0 for Windows
; SEQ ID NO 2
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide (HA residues 306-318)
US-09-768-872-2

Query Match 32.5%; Score 66; DB 10; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKYVKQNTLKLAT 13
Db 1 PKYVKQNTLKLAT 13

RESULT 11
US-09-245-487B-23
; Sequence 23, Application US/09245487B
; Patent No. US20020058787A1
; GENERAL INFORMATION:
; APPLICANT: Strominger, Jack L.
; APPLICANT: FALK, Kirsten
; APPLICANT: ROTZSCHE, Olaf
; TITLE OF INVENTION: MHC Binding Peptide Oligomers and Methods of Use
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault
; STREET: 125 High St.



CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/245,487B  
FILING DATE: 05-Feb-1999  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: TWOMEY, Michael J.  
REGISTRATION NUMBER: 38,349  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 23:  
US-09-245-487B-23

Query Match 32.5%; Score 66; DB 10; Length 13;  
Best Local Similarity 100.0%; Pred.No. 0.0016;  
Matches 13; Conservative 0; Mismatches 0; Indels

Qy      1 PKYVKQNTLKLAT 13  
         |||  
Db      1 PKYVKQNTLKLAT 13

```

RESULT 12
US-09-756-983-8
; Sequence 8, Application US/09756983
; Patent No. US2002012818A1
; GENERAL INFORMATION:
; APPLICANT: Albani, Salvatore
; TITLE OF INVENTION: METHODS FOR ISOLATION, QUANTIFICATION,
; TITLE OF INVENTION: CHARACTERIZATION AND MODULATION OF
; TITLE OF INVENTION: ANTIGEN-SPECIFIC T CELLS
; FILE REFERENCE: 246/285-CIP
; CURRENT APPLICATION NUMBER: US/09/756,983
; CURRENT FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/105,018
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: 09/421,506
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: PCT/US99/2466
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized peptide derived from the i
US-09-756-983-8

```

Query Match 32.5%; Score 66; DB 10; Length 13;  
Best Local Similarity 100.0%; Pred.No. 0.0016;  
Matches 13; Conservative 0; Mismatches 0; Indels

Qy 1 PKYVKQNTLKLAT 13  
|||  
Db 1 PKYVKQNTLKLAT 13

```

RESULT 13
US-09-756-983-14
; Sequence 14, Application US/09756983
; Patent No. US2002012818A1
; GENERAL INFORMATION:
; APPLICANT: Albani, Salvatore
; TITLE OF INVENTION: METHODS FOR ISOLATION, QUANTIFICATION,
; TITLE OF INVENTION: CHARACTERIZATION AND MODULATION OF
; TITLE OF INVENTION: ANTIGEN-SPECIFIC T CELLS
; FILE REFERENCE: 246/285-CIP
; CURRENT APPLICATION NUMBER: US/09/756, 983
; CURRENT FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/105, 018
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: 09/421, 506
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: PCT/US99/2466
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-756-983-14

```

Query Match	32.5%	Score 66;	DB 10;	Length 13;
Best Local Similarity	100.0%;	Pred. No. 0.0016;		
Matches 13;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy 1 PKYVKQNTLKLAT 13  
| | | | | | | | | |  
Db 1 PKYVKQNTLKLAT 13

```

RESULT 14
US-10-044-034-5
; Sequence 5, Application US/10044034
; Patent No. US20020169264A1
; GENERAL INFORMATION:
; APPLICANT: JACKSON, DAVID C.
; APPLICANT: O'BRIEN-SIMPSON, NEIL M.
; APPLICANT: BROWN, LORENA E.
; APPLICANT: EDE, NICHOLAS J.
; APPLICANT: BRANDT, EVELYN R.
; APPLICANT: GOOD, MICHAEL F.
; TITLE OF INVENTION: POLYMERS INCORPORATING PEPTIDES
; FILE REFERENCE: FBRC:006
; CURRENT APPLICATION NUMBER: US/10/044,034
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: P05071
; PRIOR FILING DATE: 1997-02-11
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 5
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-044-034-5

```

Query Match	31.0%	Score 63	DB 9	Length 17
Best Local Similarity	92.9%	Pred. No.	0.0056	
Matches 13	Conservative	0	Mismatches	1
			Indels	0
			Gaps	0

Qy 16 KGPKYVKQNTLKL 29  
I I I I I I I I I I  
Db 4 KYPKYVKQNTLKL 17

```
RESULT 15
US-09-987-137-7
; Sequence 7, Application US/09987137
; Patent No. US2002017554A1
; GENERAL INFORMATION:
; APPLICANT: Verheijden, Gijbertus Franciscus Maria
; APPLICANT: Boots, Anna Maria Helena
; TITLE OF INVENTION: No. US2002017554A1 Peptides for use in Treatment of T-cell Med
; TITLE OF INVENTION: Cartilage Destruction in Auto-immune Diseases
; FILE REFERENCE: Verheijdenseq
; CURRENT APPLICATION NUMBER: US/09/987,137
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: 08/981,340
; PRIOR FILING DATE: 1997-12-18
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-987-137-7

Query Match      30.5%; Score 62; DB 9; Length 13;
Best Local Similarity 92.3%; Pred. No. 0.0056;
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PKYVKQNTIKLAT 13
   ||:|||||
Db 1 PKFVKQNTIKLAT 13

Search completed: December 4, 2002, 13:05:30
Job time : 4.66667 secs
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 12:49:07 : Search time 6.66667 Seconds  
(without alignments)  
176.538 Million cell updates/sec

Title: US-09-362-731A-4  
Perfect score: 203  
Sequence: 1 PKYVKQNTLKLATGKPKYVYKQNTLKLATGKGVIGIK 40

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents.AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81.5	40.1	347	1 US-08-229-781-58	Sequence 58, Appl
2	81.5	40.1	347	1 US-08-630-918-58	Sequence 58, Appl
3	81.5	40.1	347	4 US-09-004-422-58	Sequence 58, Appl
4	80.5	39.7	566	4 US-09-232-468A-22	Sequence 22, Appl
5	77	37.9	23	4 US-08-537-928-25	Sequence 25, Appl
6	77	37.9	25	2 US-08-484-905-24	Sequence 24, Appl
7	77	37.9	25	3 US-08-481-985B-24	Sequence 24, Appl
8	77	37.9	25	4 US-08-370-476-24	Sequence 24, Appl
9	77	37.9	25	4 US-09-248-588-85	Sequence 85, Appl
10	77	37.9	25	4 US-09-082-279B-1063	Sequence 1063, Ap
11	77	37.9	25	4 US-09-315-304B-1063	Sequence 1063, Ap
12	75.5	37.2	570	2 US-08-453-848-7	Sequence 7, Appli
13	75.5	37.2	570	4 US-09-169-027-7	Sequence 7, Appli
14	75.5	37.2	571	2 US-08-453-848-15	Sequence 15, Appl
15	75.5	37.2	571	2 US-08-453-848-21	Sequence 21, Appl
16	75.5	37.2	571	4 US-09-169-027-15	Sequence 15, Appl
17	75.5	37.2	571	4 US-09-169-027-21	Sequence 21, Appl
18	74	36.5	347	4 US-09-217-293-1	Sequence 1, Appli
19	72	35.5	14	3 US-09-125-517A-61	Sequence 61, Appl
20	72	35.5	14	4 US-09-319-870A-9	Sequence 9, Appli
21	72	35.5	14	4 US-09-308-175A-21	Sequence 21, Appl
22	72	35.5	15	3 US-09-125-517A-60	Sequence 60, Appl
23	72	35.5	16	4 US-09-284-625-24	Sequence 24, Appl
24	72	35.5	16	4 US-09-284-625-26	Sequence 24, Appl
25	71	35.0	25	4 US-08-992-877-7	Sequence 7, Appli
26	70	34.5	23	4 US-08-537-928-26	Sequence 26, Appl
27	66	32.5	13	1 US-07-831-895C-5	Sequence 5, Appli

28	66	32.5	13	1	US-08-305-871A-1	Sequence 1, Appli
29	66	32.5	13	1	US-08-465-167A-22	Sequence 22, Appl
30	66	32.5	13	2	US-08-190-24	Sequence 24, Appl
31	66	32.5	13	2	US-08-596-387B-8	Sequence 8, Appli
32	66	32.5	13	2	US-08-488-379-24	Sequence 24, Appl
33	66	32.5	13	2	US-08-968-676-164	Sequence 164, App
34	66	32.5	13	4	US-09-055-475-1	Sequence 1, Appli
35	66	32.5	13	4	US-09-067-615-8	Sequence 8, Appli
36	66	32.5	13	4	US-08-464-496-18	Sequence 18, Appl
37	66	32.5	13	4	US-09-194-285-53	Sequence 53, Appl
38	66	32.5	13	4	US-08-788-822A-2	Sequence 2, Appli
39	66	32.5	13	4	US-08-197-484-99	Sequence 99, Appl
40	66	32.5	13	4	US-08-627-820-22	Sequence 22, Appl
41	66	32.5	13	5	PCT-US92-07218-18	Sequence 18, Appl
42	66	32.5	13	5	PCT-US93-07545-24	Sequence 24, Appl
43	66	32.5	13	5	PCT-US94-10257A-20	Sequence 20, Appl
44	66	32.5	13	5	PCT-US95-02121-99	Sequence 99, Appl
45	66	32.5	13	5	PCT-US95-04121-53	Sequence 53, Appl

ALIGNMENTS

RESULT 1  
US-08-229-781-58  
; Sequence 58, Application US/08229781  
; Patent No. 5589174  
; GENERAL INFORMATION:  
; APPLICANT: Yoshinobu OKUNO et al.  
; TITLE OF INVENTION: ANTI-HUMAN INFLUENZA VIRUS ANTIBODY  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/229,781  
; FILING DATE: April 19, 1994  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/054,016  
; FILING DATE: April 29, 1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek, Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX:  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 58:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 347 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHEetical:  
; ANTI-SENSE:  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; ORGANISM:  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:



QY 6 QNTLKLATGKKGPYVVKONTLKLATGKGV 35  
||| : | ||||||||| :  
Db 311 QNVNKITYG-ACPKYVVKONTLKLATGMRI 339

CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Potter, Jane E. R.  
REGISTRATION NUMBER: 33,332  
REFERENCE/DOCKET NUMBER: 03495.0106-03000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 25 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-484-905-24

Query Match 37.9%; Score 77; DB 2; Length 25;  
Best Local Similarity 83.3%; Pred. No. 0.00026;  
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 18 PKYVKONTLKLATGKGV 35  
||||| : I  
Db 2 PKYVKONTLKLATGMNV 19

RESULT 7  
US-08-481-985B-24  
; Sequence 24, Application US/08481985B  
; Patent No. 6011146  
; GENERAL INFORMATION:  
; APPLICANT: Mottez, Estelle  
; APPLICANT: Abastado, Jean-Pierre  
; APPLICANT: Kourilsky, Philippe  
; TITLE OF INVENTION: Altered Major Histocompatibility Complex  
; NUMBER OF SEQUENCES: 148  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
; STREET: 1300 I Street, N.W., Suite 700  
; CITY: Washington  
; STATE: D.C.  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/481,985B  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/801,818  
; FILING DATE: 05-DEC-1991  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/792,473  
; FILING DATE: 15-NOV-1991  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meyers, Kenneth J.  
; REGISTRATION NUMBER: 25,146  
; REFERENCE/DOCKET NUMBER: 03495.0106-04000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-408-4000  
; TELEFAX: 202-408-4400  
; INFORMATION FOR SEQ ID NO: 24:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 25 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide

US-08-481-985B-24

Query Match 37.9%; Score 77; DB 3; Length 25;  
Best Local Similarity 83.3%; Pred. No. 0.00026;  
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 18 PKYVKONTLKLATGKGV 35  
||||| : I  
Db 2 PKYVKONTLKLATGMNV 19

RESULT 8  
US-08-370-476-24  
; Sequence 24, Application US/08370476  
; Patent No. 6153408  
; GENERAL INFORMATION:  
; APPLICANT: Mottez, Estelle  
; APPLICANT: Abastado, Jean-Pierre  
; APPLICANT: Kourilsky, Philippe  
; APPLICANT: Lone, Yu-Chun  
; APPLICANT: Ojcius, David  
; APPLICANT: Casrouge, Armanda  
; TITLE OF INVENTION: Altered Major Histocompatibility Complex  
; NUMBER OF SEQUENCES: 127  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
; STREET: 1300 I Street, N.W., Suite 700  
; CITY: Washington  
; STATE: D.C.  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/370,476  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/117,575  
; FILING DATE: 07-SEP-1993  
; APPLICATION NUMBER: US 08/072,787  
; FILING DATE: 06-JUN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/801,818  
; FILING DATE: 05-DEC-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/792,473  
; FILING DATE: 15-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meyers, Kenneth J.  
; REGISTRATION NUMBER: 25,146  
; REFERENCE/DOCKET NUMBER: 05243.0001-01000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-408-4000  
; TELEFAX: 202-408-4400  
; INFORMATION FOR SEQ ID NO: 24:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 25 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-370-476-24

Query Match 37.9%; Score 77; DB 4; Length 25;  
Best Local Similarity 83.3%; Pred. No. 0.00026;  
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 18 PKYVKONTLKLATGKGV 35  
||||| : I

Db 2 PKYVKQNTLKLATGMRNV 19

## RESULT 9

US-09-248-588-85  
; Sequence 85, Application US/09248588  
; Patent No. 6231864

; GENERAL INFORMATION:  
; APPLICANT: Birkett, Ashley J.

; TITLE OF INVENTION: Strategically Modified Hepatitis B Core Proteins and  
; TITLE OF INVENTION: their Derivatives

; FILE REFERENCE: SYN-101 4564/69529

; CURRENT APPLICATION NUMBER: US/09/248,588  
; CURRENT FILING DATE: 1999-02-11

; EARLIER APPLICATION NUMBER: 60/074537  
; EARLIER FILING DATE: 1998-02-12

; NUMBER OF SEQ ID NOS: 113  
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 85

; LENGTH: 25

; TYPE: PRT

; ORGANISM: Influenza virus

US-09-248-588-85

Query Match 37.9%; Score 77; DB 4; Length 25;

Best Local Similarity 83.3%; Pred. No. 0.00026;

Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 18 PKYVKQNTLKLATGKGV 35

|||||

Db 2 PKYVKQNTLKLATGMRNV 19

## RESULT 10

US-09-082-279B-1063

; Sequence 1063, Application US/09082279B

; Patent No. 6258782

; GENERAL INFORMATION:

; APPLICANT: Barney, Shawn

; APPLICANT: Guthrie, Kelly

; APPLICANT: Merutka, Gene

; APPLICANT: Anwer, Mohmed

; APPLICANT: Lambert, Dennis

; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED

; TITLE OF INVENTION: PHARMACOKINETIC PROPERTIES

; FILE REFERENCE: 7872-043

; CURRENT APPLICATION NUMBER: US/09/082,279B

; CURRENT FILING DATE: 1998-05-20

; NUMBER OF SEQ ID NOS: 1515

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1063

; LENGTH: 25

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Core polypeptide

US-09-082-279B-1063

Query Match 37.9%; Score 77; DB 4; Length 25;

Best Local Similarity 83.3%; Pred. No. 0.00026;

Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 18 PKYVKQNTLKLATGKGV 35

|||||

Db 2 PKYVKQNTLKLATGMRNV 19

## RESULT 11

US-09-315-304B-1063

; Sequence 1063, Application US/09315304B

; Patent No. 6348568

; GENERAL INFORMATION:

; APPLICANT: Barney, S.

; APPLICANT: Guthrie, K.

; APPLICANT: Merutka, G.

; APPLICANT: Anwer, M.

; APPLICANT: Lambert, D.

; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC

; FILE REFERENCE: 7872-052

; CURRENT APPLICATION NUMBER: US/09/315,304B

; CURRENT FILING DATE: 1999-05-20

; NUMBER OF SEQ ID NOS: 1667

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1063

; LENGTH: 25

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Core polypeptide

US-09-315-304B-1063

Query Match 37.9%; Score 77; DB 4; Length 25;

Best Local Similarity 83.3%; Pred. No. 0.00026;

Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 18 PKYVKQNTLKLATGKGV 35

|||||

Db 2 PKYVKQNTLKLATGMRNV 19

RESULT 12

US-08-453-848-7

; Sequence 7, Application US/08453848

; Patent No. 5858368

; GENERAL INFORMATION:

; APPLICANT: Smith, Gale Eugene

; APPLICANT: Volvovitz, Franklin

; APPLICANT: Wilkinson, Bethanie Eident

; APPLICANT: Voznesensky, Andrei I.

; APPLICANT: Hackett, Craig Stanway

; TITLE OF INVENTION: A METHOD FOR PRODUCING INFLUENZA

; TITLE OF INVENTION: HEMAGGLUTININ MULTIVALENT VACCINES

; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Patrea L. Pabst

; STREET: 2800 One Atlantic Center

; STREET: 1201 West Peachtree Street

; CITY: Atlanta

; STATE: GA

; COUNTRY: USA

; ZIP: 30309-3450

; COMPUTER READABLE FORM:

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/453,848

; FILING DATE: 30-MAY-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/120,607

; FILING DATE: 13-SEPT-1993

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Pabst, Patrea L.

; REGISTRATION NUMBER: 31,284

; REFERENCE/DOCKET NUMBER: MGS101CIP

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (404)-873-8794

; TELEFAX: (404)-873-8795

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

LENGTH: 570 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Influenza virus  
INDIVIDUAL ISOLATE: A/Beijing/32/92 rHA  
FEATURE:  
NAME/KEY: AcNPV 61K protein signal sequence  
LOCATION: 1 to 18  
FEATURE:  
NAME/KEY: mature rHA  
LOCATION: 19 to 552  
US-08-453-848-7

Query Match 37.2%; Score 75.5; DB 2; Length 570;  
Best Local Similarity 56.7%; Pred. No. 0.016;  
Matches 17; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

QY 6 QNTLKLATGKKGPKYVKQNTLKLATGKGV 35  
|| : : | | : ||||| : |  
DB 315 QNVNRTYTG-ACPRYVKQNTLKLATGMNV 343

RESULT 13  
US-09-169-027-7  
Sequence 7, Application US/09169027  
Patent No. 6245532  
GENERAL INFORMATION:  
APPLICANT: Smith, Gale Eugene  
APPLICANT: Volvovitz, Franklin  
APPLICANT: Wilkinson, Bethanie Eident  
APPLICANT: Voznesensky, Andrei I.  
APPLICANT: Hackett, Craig Stanway  
TITLE OF INVENTION: A METHOD FOR PRODUCING INFLUENZA  
TITLE OF INVENTION: HEMAGGLUTININ MULTIVALENT VACCINES  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Patrea L. Pabst  
STREET: 2800 One Atlantic Center  
STREET: 1201 West Peachtree Street  
CITY: Atlanta  
STATE: GA  
COUNTRY: USA  
ZIP: 30309-3450  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/169,027  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/453,848  
FILING DATE: 30-MAY-1995  
APPLICATION NUMBER: 08/120,607  
FILING DATE: 13-SEPT-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: MGS101CIP  
TELEPHONE: (404)-873-8794  
TELEFAX: (404)-873-8795  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 570 amino acids

TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Influenza virus  
INDIVIDUAL ISOLATE: A/Beijing/32/92 rHA  
FEATURE:  
NAME/KEY: AcNPV 61K protein signal sequence  
LOCATION: 1 to 18  
FEATURE:  
NAME/KEY: mature rHA  
LOCATION: 19 to 552  
US-09-169-027-7

Query Match 37.2%; Score 75.5; DB 4; Length 570;  
Best Local Similarity 56.7%; Pred. No. 0.016;  
Matches 17; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

QY 6 QNTLKLATGKKGPKYVKQNTLKLATGKGV 35  
|| : : | | : ||||| : |  
DB 315 QNVNRTYTG-ACPRYVKQNTLKLATGMNV 343

RESULT 14  
US-08-453-848-15  
Sequence 15, Application US/08453848  
Patent No. 5858368  
GENERAL INFORMATION:  
APPLICANT: Smith, Gale Eugene  
APPLICANT: Volvovitz, Franklin  
APPLICANT: Wilkinson, Bethanie Eident  
APPLICANT: Voznesensky, Andrei I.  
APPLICANT: Hackett, Craig Stanway  
TITLE OF INVENTION: A METHOD FOR PRODUCING INFLUENZA  
TITLE OF INVENTION: HEMAGGLUTININ MULTIVALENT VACCINES  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Patrea L. Pabst  
STREET: 2800 One Atlantic Center  
STREET: 1201 West Peachtree Street  
CITY: Atlanta  
STATE: GA  
COUNTRY: USA  
ZIP: 30309-3450  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/453,848  
FILING DATE: 30-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/120,607  
FILING DATE: 13-SEPT-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: MGS101CIP  
TELEPHONE: (404)-873-8794  
TELEFAX: (404)-873-8795  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 571 amino acids  
TYPE: amino acid  
STRANDEDNESS: single



```

; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Influenza virus
; INDIVIDUAL ISOLATE: A/Shandong/9/93 rHA
; FEATURE:
; NAME/KEY: ACNPV 61K protein signal sequence
; LOCATION: 1 to 18
; FEATURE:
; NAME/KEY: mature rHA
; LOCATION: 19 to 553
;
US-08-453-848-15

Query Match          37.2%; Score 75.5; DB 2; Length 571;
Best Local Similarity 56.7%; Pred. No. 0.016;
Matches 17; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

QY 6 QNTLKLATGKKGPKYVKQNTLKLATGKKG 35
Db 316 QNVNRITYG-ACPRYVKQNTLKLATGMRNV 344

RESULT 15
US-08-453-848-21
; Sequence 21, Application US/08453848
; Patent No. 5858368
; GENERAL INFORMATION:
; APPLICANT: Smith, Gale Eugene
; APPLICANT: Volvovitz, Franklin
; APPLICANT: Wilkinson, Bethanie Eident
; APPLICANT: Voznesensky, Andrei I.
; APPLICANT: Hackett, Craig Stanway
; TITLE OF INVENTION: A METHOD FOR PRODUCING INFLUENZA
; TITLE OF INVENTION: HEMAGGLUTININ MULTIVALENT VACCINES
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: GA
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,848
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/120,607
; FILING DATE: 13-SEPT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: MGS101CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)-873-8794
; TELEFAX: (404)-873-8795
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 571 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
```

```

; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Influenza virus
; INDIVIDUAL ISOLATE: A/Johannesburg/33/94 rHA
; FEATURE:
; NAME/KEY: ACNPV 61K protein signal sequence
; LOCATION: 1 to 18
; FEATURE:
; NAME/KEY: mature rHA
; LOCATION: 19 to 569
;
US-08-453-848-21

Query Match          37.2%; Score 75.5; DB 2; Length 571;
Best Local Similarity 56.7%; Pred. No. 0.016;
Matches 17; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

QY 6 QNTLKLATGKKGPKYVKQNTLKLATGKKG 35
Db 316 QNVNRITYG-ACPRYVKQNTLKLATGMRNV 344

Search completed: December 4, 2002, 12:55:09
Job time : 8.66667 secs
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 12:49:07 : Search time 6.26667 Seconds  
(without alignments)  
490.899 Million cell updates/sec

Title: US-09-362-731a-5

Perfect score: 181

Sequence: 1 QYIKANSKFTGITELGGCHGSEPCNIHRGKPF 32

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_73:\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	92	50.8	146	2 A60381	allergen Der p II
2	89	49.2	129	2 A61501	allergen Der f II
3	89	49.2	129	2 JU0394	allergen Der f II
4	89	49.2	138	2 B61241	allergen Der f II
5	89	49.2	138	2 A61241	allergen Der f II
6	74	40.9	1315	1 BTCLFN	tentoxilysin (EC 3
7	53.5	29.6	403	2 T02513	hypothetical prote
8	52	28.7	289	2 A45369	insect-selective n
9	52	28.7	1884	2 JC4975	plexin 2 precursor
10	51	28.2	307	2 A71807	hypothetical prote
11	50.5	27.9	244	2 S29982	class II histocomp
12	49.5	27.3	253	2 T30381	hypothetical prote
13	49.5	27.3	329	2 D85956	hypothetical prote
14	49.5	27.3	329	2 A91111	hypothetical prote
15	49	27.1	170	2 T28140	natural killer cel
16	49	27.1	238	2 I48605	insulin-like growt
17	49	27.1	510	2 D96741	hypothetical prote
18	48.5	26.8	378	2 S69683	hypothetical prote
19	48	26.5	484	2 A71456	hypothetical prote
20	48	26.5	484	2 B75016	hypothetical prote
21	48	26.5	502	2 T05135	hypothetical prote
22	48	26.5	532	2 T52442	hypothetical prote
23	48	26.5	612	2 I64241	glucose inhibited
24	48	26.5	817	2 T24063	hypothetical prote
25	47.5	26.2	173	2 T00970	hypothetical prote
26	47.5	26.2	533	2 A47143	retinal pigment mi
27	47	26.0	38	2 G64583	hypothetical prote
28	47	26.0	158	2 G82945	hypothetical prote
29	47	26.0	398	2 T09982	rfe protein - Myco

30 47 26.0 399 2 C87051  
31 47 26.0 404 2 B70774  
32 47 26.0 630 2 S41314  
33 46.5 25.7 141 2 S66499  
34 46.5 25.7 245 2 S29980  
35 46.5 25.7 285 2 C97279  
36 46.5 25.7 358 2 D64113  
37 46.5 25.7 379 2 A81140  
38 46.5 25.7 379 2 C81880  
39 46 25.4 191 2 T29502  
40 46 25.4 264 2 T45255  
41 46 25.4 265 2 H87067  
42 46 25.4 275 2 S77462  
43 46 25.4 289 2 T18919  
44 46 25.4 339 2 C42276  
45 46 25.4 379 2 T44656

probable glycosylt  
probable rfe prote  
hypothetical prote  
allergen Lep d I.0  
class II histocomp  
thioredoxin reduct  
homoserine O-acety  
homoserine O-acety  
probable homoserin  
hypothetical prote  
probable membrane  
probable conserved  
hypothetical prote  
hypothetical prote  
O-methyltransferas  
homoserine O-acety

#### ALIGNMENTS

##### RESULT 1

A60381  
allergen Der p II precursor - house-dust mite (Dermatophagoides pteronyssinus)  
C:Species: Dermatophagoides pteronyssinus  
C>Date: 03-Mar-1993 #sequence\_revision 03-Mar-1993 #text\_change 13-Sep-1998  
C:Accession: A60381  
R:Chua, K.Y.; Doyle, C.R.; Simpson, R.J.; Turner, K.J.; Stewart, G.A.; Thomas, W.R.  
Int. Arch. Allergy Appl. Immunol. 91, 118-123, 1990  
A:Title: Isolation of cDNA coding for the major mite allergen Der p II by IgE plaque  
A:Reference number: A60381; MUID:90256301; PMID:2341191  
A:Accession: A60381  
A>Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-146 <CHU>  
C:Superfamily: allergen Der p II  
F:1-17/Domain: signal sequence #status predicted <SIG>  
F:18-146/Product: allergen Der p II #status predicted <MAT>

Query Match 50.8%; Score 92; DB 2; Length 146;  
Best Local Similarity 93.8%; Pred. No. 1.4e-05;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 17 GCHGSEPCNIHRGKPF 32  
||||||| |||||  
DB 37 GCHGSEPCNIHRGKPF 52

##### RESULT 2

A61501  
allergen Der f II - house-dust mite (Dermatophagoides farinae) (fragment)  
C:Species: Dermatophagoides farinae  
C>Date: 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 13-Sep-1998  
C:Accession: A61501  
R:Frudinger, M.; Chua, K.Y.; Thomas, W.R.  
Clin. Exp. Allergy 21, 33-37, 1991  
A:Title: cDNA encoding the major mite allergen Der f II.  
A:Reference number: A61501; MUID:91215495; PMID:2021876  
A:Accession: A61501  
A>Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-129 <TRU>  
C:Superfamily: allergen Der p II

Query Match 49.2%; Score 89; DB 2; Length 129;  
Best Local Similarity 87.5%; Pred. No. 3.4e-05;  
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 GCHGSEPCNIHRGKPF 32  
||||||| |||||  
DB 20 GCHGSDPCNIHRGKPF 35

```

RESULT 3
JU0394
allergen Der f II (pFL2) - house-dust mite (Dermatophagoides farinae)
C:Species: Dermatophagoides farinae
C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 17-Mar-1999
C:Accession: JU0394
R:Yuuki, T.; Okumura, Y.; Ando, T.; Yamakawa, H.; Suko, M.; Haida, M.; Okudaira, H.
Agric. Biol. Chem. 55, 1233-1238, 1991
A:Title: Cloning and expression of cDNA coding for the major house dust mite allergen Der
A:Reference number: PS0417; MUID:91291341; PMID:1368682
A:Accession: JU0394
A:Molecule type: mRNA
A:Residues: 1-129 <YUU>
C:Superfamily: allergen Der p II

Query Match 49.2%; Score 89; DB 2; Length 129;
Best Local Similarity 87.5%; Pred. No. 3.4e-05;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 17 GCHGSEPCNIHRGKPF 32
|||||:|||||
Db 20 GCHGSDPCIHRGKPF 35

RESULT 4
B61241
allergen Der f II precursor - house-dust mite (Dermatophagoides farinae) (fragment)
C:Species: Dermatophagoides farinae
C>Date: 12-May-1994 #sequence_revision 27-Jun-1994 #text_change 13-Sep-1998
C:Accession: B61241; JU0395
R:Yuuki, T.; Okumura, Y.; Ando, T.; Yamakawa, H.; Suko, M.; Haida, M.; Okudaira
Int. Arch. Allergy Appl. Immunol. 94, 354-356, 1991
A:Title: Synthesis of biologically active recombinant Der f II.
A:Reference number: A61241; MUID:92040281; PMID:1937898
A:Accession: B61241
A:Molecule type: mRNA
A:Residues: 1-138 <YUU>
C:Superfamily: allergen Der p II
F:1-9/Domain: signal sequence (fragment) #status predicted <SIG>
F:10-138/Product: allergen Der f II #status predicted <MAT>

Query Match 49.2%; Score 89; DB 2; Length 138;
Best Local Similarity 87.5%; Pred. No. 3.6e-05;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 17 GCHGSEPCNIHRGKPF 32
|||||:|||||
Db 29 GCHGSDPCIHRGKPF 44

RESULT 5
A61241
allergen Der f II precursor - house-dust mite (Dermatophagoides farinae) (fragment)
C:Species: Dermatophagoides farinae
C>Date: 12-May-1994 #sequence_revision 27-Jun-1994 #text_change 13-Sep-1998
C:Accession: A61241; PS0417
R:Yuuki, T.; Okumura, Y.; Ando, T.; Yamakawa, H.; Suko, M.; Haida, M.; Dohi, M.; Okudaira
Int. Arch. Allergy Appl. Immunol. 94, 354-356, 1991
A:Title: Synthesis of biologically active recombinant Der f II.
A:Reference number: A61241; MUID:92040281; PMID:1937898
A:Accession: A61241
A:Molecule type: mRNA
A:Residues: 1-138 <YUU>
A:Note: part of this sequence, including the amino end of the mature protein, was confir
C:Superfamily: allergen Der p II
F:1-9/Domain: signal sequence (fragment) #status predicted <SIG>
F:10-138/Product: allergen Der f II #status experimental <MAT>

Query Match 49.2%; Score 89; DB 2; Length 138;
Best Local Similarity 87.5%; Pred. No. 3.6e-05;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 17 GCHGSEPCNIHRGKPF 32
|||||:|||||
Db 29 GCHGSDPCIHRGKPF 44

```

```

|||||:|||||
Db 29 GCHGSDPCIHRGKPF 44

RESULT 6
B7CLTN
tentoxilysin (EC 3.4.24.68) precursor - Clostridium tetani
N:Alternate names: tetanus neurotoxin
C:Species: Clostridium tetani
C>Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 03-Jun-2002
C:Accession: A25689; A25757; A25194; B25194; A60759; S69348; S09364
R:Eisel, U.; Jarausch, W.; Goretzki, K.; Henschen, A.; Engels, J.; Weller, U.; Hudel,
EMBO J. 5, 2495-2502, 1986
A:Title: Tetanus toxin: primary structure, expression in E. coli, and homology with b
A:Reference number: A25689; MUID:87053814; PMID:3536478
A:Accession: A25689
A:Molecule type: DNA
A:Residues: 1-1315 <EIS>
A:Cross-references: GB:X04436; NID:940769; PIDN:CAA28033.1; PID:g40770
R:Fairweather, N.F.; Lyness, V.A.
Nucleic Acids Res. 14, 7809-7812, 1986
A:Title: The complete nucleotide sequence of tetanus toxin.
A:Reference number: A25757; MUID:87040747; PMID:3774547
A:Accession: A25757
A:Molecule type: DNA
A:Residues: 1-1315 <FAI>
A:Cross-references: GB:X06214; NID:g40773; PIDN:CAA29564.1; PID:g40774
A:Experimental source: strain CN3911
R:Fairweather, N.F.; Lyness, V.A.; Pickard, D.J.; Allen, G.; Thomson, R.O.
J. Bacteriol. 165, 21-27, 1986
A:Title: Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C i
A:Reference number: A25194; MUID:86085672; PMID:3510187
A:Accession: A25194
A:Molecule type: DNA
A:Residues: 743-1315 <FA2>
A:Cross-references: GB:M12739; NID:g144920; PIDN:AAA23282.1; PID:g144921
A:Accession: B25194
A:Molecule type: protein
A:Residues: 865-894 <FA3>
R:Matsuda, M.; Lei, D.L.; Sugimoto, N.; Ozutsumi, K.; Okabe, T.
Infect. Immun. 57, 3588-3593, 1989
A:Title: Isolation, purification, and characterization of fragment B, the NH-2-termin
A:Reference number: A60759; MUID:90035436; PMID:2478476
A:Accession: A60759
A:Molecule type: protein
A:Residues: 461-475 <MAT>
R:Demotz, S.; Lanzavecchia, L.; Eisel, U.; Niemann, H.; Widmann, C.; Corradin, G.
J. Immunol. 142, 394-402, 1989
A:Title: Delineation of several DR-restricted tetanus toxin T cell epitopes.
A:Reference number: JS0098; MUID:89093918; PMID:2463305
A:Contents: annotation; epitope region
R:Schiavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; DasGupta,
Nature 359, 832-835, 1992
A:Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteo
A:Reference number: S27125; MUID:93063293; PMID:1331807
A:Contents: annotation
R:de Filippis, V.; Vangelista, L.; Schiavo, G.; Tonello, F.; Montecucco, C.
Eur. J. Biochem. 229, 61-69, 1995
A:Title: Structural studies on the zinc-endopeptidase light chain of tetanus neurotox
A:Reference number: S69348; MUID:95262688; PMID:7744050
A:Accession: S69348
A:Molecule type: protein
A:Residues: 2-31 <DEF>
C:Comment: The source of this protein was an extrachromosomal plasmid.
C:Comment: The precursor is cleaved by endogenous proteinase activity to form light (
dual chains are not toxic when separated). The amino end of the heavy chain (fragment
C:Comment: Fragment B forms ion channels in a lipid bilayer. Fragment C binds to gang
C:Comment: This potent neurotoxin binds to peripheral neuronal synapses, is internali
presynaptic neurons. It inhibits neurotransmitter release by proteolytic cleavage of
C:Function:
A:Description: blocks neuroexocytosis via hydrolysis of a Gln-Phe peptide bond in syn
C:Superfamily: tetanus toxin
C:Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc

```

F:2-457/Product: tentoxylisin light chain (fragment A) #status predicted <TTL>  
F:461-1315/Product: tentoxylisin heavy chain (fragment B.C) #status experimental <TTX>  
F:461-864/Domain: channel forming (fragment B) #status predicted <TTX>  
F:865-1315/Domain: ganglioside binding (fragment C) #status predicted <TXC>  
F:233,237/Binding site: zinc (His) #status predicted  
F:234/Active site: Glu #status predicted

Query Match 40.9%; Score 74; DB 1; Length 1315;

Best Local Similarity 100.0%; Pred. No. 0.037;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15

Db 830 QYIKANSKFIGITEL 844

#### RESULT 7

T02513

hypothetical protein At2g38320 [imported] - Arabidopsis thaliana

N:Alternate names: hypothetical protein T19C21.19

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 16-Feb-2001

R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul

submitted to the EMBL Data Library, August 1998

A:Description: Arabidopsis thaliana chromosome II BAC T19C21 genomic sequence.

A:Reference number: Z14676

A:Accession: T02513

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-403 <ROU>

A:Cross-references: EMBL:AC004683; NID:g3395421; PID:g3395440

A:Experimental source: cultivar Columbia

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: E84803

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-403 <STO>

A:Cross-references: GB:AE002093; NID:g3395440; PIDN:AA28772.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g38320; T19C21.19

A:Map position: 2

A:Introns: 107/2; 166/3; 232/2; 285/2

#### Query Match

Best Local Similarity 29.6%; Score 53.5; DB 2; Length 403;

Matches 12; Conservative 4; Mismatches 10; Indels 1; Gaps 1;

Qy 4 KANSKFIGITELGGCHSEPCNIHRGK 30

Db 33 KENPQSHGVTRDGGDSGRE-CNLFEKG 58

#### RESULT 8

A45369

insect-selective neurotoxin TxP-I homolog - straw itch mite

C:Species: Pyemotes tritici (straw itch mite)

C:Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 22-Oct-1999

R:Tomalski, M.D.; Hutchinson, K.; Todd, J.; Miller, L.K.

Toxicol 31, 319-326, 1993

A:Title: Identification and characterization of tox21A: a mite cDNA encoding a paralytic

A:Reference number: A45369; MUID:93227269; PMID:8470135

A:Accession: A45369

A:Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-289 <TOM>

A:Cross-references: GB:S58065; NID:g299121; PID:g299122

A:Note: sequence extracted from NCBI backbone (NCBIN:129517, NCBIP:129520)  
C:Superfamily: straw itch mite toxin TxP-I

Query Match 28.7%; Score 52; DB 2; Length 289;

Best Local Similarity 41.9%; Pred. No. 11;

Matches 13; Conservative 4; Mismatches 10; Indels 4; Gaps 2;

Qy 4 KANSKFP--IGITELGGCHGSEPCNIHRGKPF 32

Db 187 KNSKYIEIVGVCLGRCHGTG--NFSNGEVF 215

#### RESULT 9

JC4975

plexin 2 precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 31-Dec-1996 #sequence\_revision 31-Dec-1996 #text\_change 05-Nov-1999

C:Accession: JC4975

R:Kameyama, T.; Murakami, Y.; Suto, F.; Kawakami, A.; Takagi, S.; Hirata, T.; Fujisaw

Biochem. Biophys. Res. Commun. 226, 396-402, 1996

A:Title: Identification of plexin family molecules in mice.

A:Reference number: JC4975; MUID:96400270; PMID:8806646

A:Accession: JC4975

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-1884 <KAM>

A:Cross-references: DBJ:086949; NID:g1655431; PIDN:BAAL3189.1; PID:d1013878; PID:g16

C:Comment: This protein is a membrane protein with cell adhesion properties.

C:Keywords: duplication; transmembrane protein

F:1-21/Domain: signal sequence #status predicted <SIG>

F:501-549/Region: cysteine-rich duplication

F:646-691/Region: cysteine-rich duplication

F:794-845/Region: cysteine-rich duplication

F:1223-1250/Domain: transmembrane #status predicted <TMM>

#### Query Match

Best Local Similarity 28.7%; Score 52; DB 2; Length 1884;

Matches 11; Conservative 4; Mismatches 10; Indels 4; Gaps 1;

Qy 3 IKANSKFIGITELGGCHGSEPCNIHRGK 31

Db 805 LKADHKF----ECGWCSEGERCTLHQHCP 829

#### RESULT 10

A71807

hypothetical protein jhp1433 - Helicobacter pylori (strain J99)

C:Species: Helicobacter pylori

A:Variety: strain J99

C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 08-Oct-1999

C:Accession: A71807

R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.

Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.

Nature 397, 176-180, 1999

A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric p

A:Reference number: A71800; MUID:99120557; PMID:9923682

A:Accession: A71807

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-307 <ARN>

A:Cross-references: GB:AE001566; GB:AE001439; NID:g4156051; PIDN:AAD07009.1; PID:g415

A:Experimental source: strain J99

C:Genetics:

A:Gene: jhp1433

C:Superfamily: conserved hypothetical protein HI0072

#### Query Match

Best Local Similarity 28.2%; Score 51; DB 2; Length 307;

Matches 8; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITELGGCHGSEPCNIHRG 29

Db 163 EIVIAKKKALGVLDIKACAGHTPFNTYKG 191

```

Best Local Similarity 48.0%; Pred. No. 29;
Matches 12; Conservative 2; Mismatches 6; Indels 5; Gaps 2;

QY 3 IKANSKFIGITELG--CGHSEPCN 25
Db 194 LKAGSLF---TETGKTGCHNISPCS 215
      ||| | | || | ||| |||
RESULT 14
A:Title: hypothetical protein ECS3857 [imported] - Escherichia coli (strain O157:H7, substrain
A:Reference number: S29982
A:Accession: S29982
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-244 <HOR>
A:Cross-references: EMBL:X70166; NID:g64369; PID:g64370
C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match 27.9%; Score 50.5; DB 2; Length 244;
Best Local Similarity 57.9%; Pred. No. 16;
Matches 11; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

QY 1 QYIKANS---KFIGITELG 16
Db 51 EYIRFNSTVGKRVGYTELG 69
      ||| | | ||| | ||| |||
RESULT 12
T30381
hypothetical protein ORF34 - Lymnaea dispar nuclear polyhedrosis virus
C:Species: Lymnaea dispar nuclear polyhedrosis virus, LdMNPV
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T30381
R:Kuzio, J.; Pearson, M.N.; Harwood, S.H.; Funk, C.J.; Evans, J.T.; Slavicek, J.M.; Rohlf
Virology 253, 17-34, 1999
A:Title: Sequence and analysis of the genome of a baculovirus pathogenic for Lymnaea d
A:Reference number: Z20836; MUID:99124785; PMID:9887315
A:Accession: T30381
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-253 <KUZ>
A:Cross-references: EMBL:AF081810; PIDN:AAC70219.1

Query Match 27.3%; Score 49.5; DB 2; Length 253;
Best Local Similarity 52.6%; Pred. No. 23;
Matches 10; Conservative 2; Mismatches 4; Indels 3; Gaps 1;

QY 14 ELGGCHGSEP---CNIHRG 29
Db 151 DLGGPGRPVVTCNVHGG 169
      ||| | | ||| | ||| |
RESULT 13
D85956
hypothetical protein Z4328 [imported] - Escherichia coli (strain O157:H7, substrain EDL9
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: D85956
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: D85956
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-329 <STO>
A:Cross-references: GB:AF005174; NID:q12517536; PIDN:AAG58112.1; GSPDB:GN00145; UWGP:Z43
A:Experimental source: strain O157:H7, substrain EDL933
A:Genetics:
A:Gene: Z4328

Query Match 27.3%; Score 49.5; DB 2; Length 329;

```

```

Best Local Similarity 48.0%; Pred. No. 29;
Matches 12; Conservative 2; Mismatches 6; Indels 5; Gaps 2;

QY 3 IKANSKFIGITELG--CGHSEPCN 25
Db 194 LKAGSLF---TETGKTGCHNISPCS 215
      ||| | | || | ||| |||
RESULT 14
A:Title: hypothetical protein ECS3857 [imported] - Escherichia coli (strain O157:H7, substrain
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: A91111
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-329 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA037280.1; PID:g13363329; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECS3857

Query Match 27.3%; Score 49.5; DB 2; Length 329;
Best Local Similarity 48.0%; Pred. No. 29;
Matches 12; Conservative 2; Mismatches 6; Indels 5; Gaps 2;

QY 3 IKANSKFIGITELG--CGHSEPCN 25
Db 194 LKAGSLF---TETGKTGCHNISPCS 215
      ||| | | || | ||| |||
RESULT 15
T28140
natural killer cell receptor homolog - chicken (fragment)
C:Species: Gallus gallus (chicken)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T28140
R:Milne, S.; Kaufman, J.; Beck, S.
submitted to the EMBL Data Library, May 1998
A:Description: DNA sequencing and analysis of the chicken major histocompatibility co
A:Reference number: Z20475
A:Accession: T28140
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-170 <MIL>
A:Cross-references: EMBL:AL023516; NID:el292539; PID:el292544; PIDN:CAA18960.1
A:Experimental source: clone CB12
C:Genetics:
A:Gene: B-NK
A:Map position: 16
A:Note: Intron positions not resolved (incomplete sequence)

Query Match 27.1%; Score 49; DB 2; Length 170;
Best Local Similarity 40.0%; Pred. No. 19;
Matches 8; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 12 ITELGGCHGSEPCNIHRGKP 31
Db 10 LTKPGGCSASTPCSMEPAGP 29
      ||| | | ||| | ||| |

```

Search completed: December 4, 2002, 12:54:25  
Job time : 8.26667 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 12:49:06 : Search time 3.46667 Seconds  
(without alignments)  
382.858 Million cell updates/sec

Title: US-09-362-731A-5  
Perfect score: 181  
Sequence: 1 QYIKANSKFTIGITELGCGHSEPCNIHRGKPF 32

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	92	50.8	146	1 ALL2_DERPT	P49278 dermatophag
2	89	49.2	146	1 ALL2_DERFA	Q00855 dermatophag
3	74	40.9	1314	1 TETX_CLOFE	P04958 clostridium
4	69	38.1	145	1 ALL2_EURMA	Q9tz22 euroglyphus
5	51	28.2	284	1 PPNK_HELPJ	Q9z181 helicobacte
6	49	27.1	238	1 IBP6_MOUSE	P47880 mus musculu
7	49	27.1	259	1 LPXA_CHRVI	Q46481 chromatium
8	48.5	26.8	378	1 URH1_YEAST	Q04179 saccharomyc
9	48	26.5	612	1 GIDA_MYCGE	P47619 mycoplasma
10	47	26.0	158	1 Y008_UREPA	Q9prd7 ureaplasma
11	47	26.0	398	1 RFE_MYCLE	P45830 mycobacteri
12	47	26.0	404	1 RPE_MYCTU	Q10606 mycobacteri
13	47	26.0	406	1 BHMT_MYCTU	Q93088 homo sapien
14	47	26.0	407	1 BHMT_MOUSE	Q33490 mus musculu
15	47	26.0	407	1 BHMT_RAT	Q09171 rattus norv
16	46.5	25.7	141	1 ALL2_TYRPU	O02380 tyrophagus
17	46.5	25.7	357	1 METX_HAEIN	P45131 haemophilus
18	46.5	25.7	379	1 METX_NEIMA	Q9jut9 neisseria m
19	46.5	25.7	379	1 METX_NEIMB	Q9zcd5 neisseria m
20	46	25.4	339	1 TCMO_STRGA	P39896 streptomyce
21	46	25.4	379	1 TCMO_LEPME	P94891 leptospira
22	45.5	25.1	141	1 ALL1_LEPDS	P80384 lepidoglyph
23	45.5	25.1	182	1 PGRP_TRINI	Q76537 trichoplusi
24	45.5	25.1	470	1 NRAM_IARUE	P31510 influenza a
25	45.5	25.1	470	1 NRAM_IATRA	P03472 influenza a
26	45.5	25.1	470	1 NRAM_IATWA	P03603 influenza a
27	45.5	25.1	4393	1 PGBM_HUMAN	P98160 homo sapien
28	45	24.9	302	1 SPRC_MOUSE	P07214 mus musculu
29	45	24.9	302	1 SPRC_MOUSE	P16975 rattus norv
30	45	24.9	382	1 METX_MYCLE	Q38874 mycobacteri
31	45	24.9	502	1 COAT_BPT6	P04535 bacterioph
32	45	24.9	521	1 COAT_BPAR1	Q9zxio bacterioph
33	45	24.9	521	1 COAT_BPT4	P04535 bacterioph

34	45	24.9	584	1	PIA3_MOUSE	O54714 mus musculu
35	45	24.9	773	1	DPOL_THEGO	P56689 thermococcu
36	45	24.9	780	1	Z084_XENLA	P18753 xenopus lae
37	44.5	24.6	125	1	LEFE_NPVOP	Q05124 orgyia pseu
38	44.5	24.6	380	1	METX_THETH	Q9ras1 thermus the
39	44.5	24.6	397	1	TDG_MOUSE	P56581 mus musculu
40	44.5	24.6	2156	1	RP1_HUMAN	P56715 homo sapien
41	44.5	24.6	2214	1	POLG_CXA24	P36290 c genome po
42	44.5	24.6	2319	1	AKA6_HUMAN	Q13023 homo sapien
43	44.5	24.6	3707	1	PGBM_MOUSE	Q05793 mus musculu
44	44	24.3	66	1	VG84_BPML5	Q05301 mycobacteri
45	44	24.3	195	1	GYRA_FIBSU	P35810 fibrobacter

#### ALIGNMENTS

RESULT 1						
ALL2_DERPT						
ID	ALL2_DERPT	STANDARD;	PRT;	146 AA.		
AC	P49278;					
DT	01-FEB-1996 (Rel. 33, Created)					
DT	01-FEB-1996 (Rel. 33, Last sequence update)					
DT	16-OCT-2001 (Rel. 40, Last annotation update)					
DE	Mite group 2 allergen Der p 2 precursor (Der p II) (DPX).					
GN	DERP2.					
OS	Dermatophagoides pteronyssinus (House-dust mite).					
OC	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;					
OC	Acariformes; Sarcophtiformes; Astigmata; Analgoidea; Pyroglyphidae;					
OC	Dermatophagoides.					
OX	NCBI_TaxID=6956;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=90256301; PubMed=2341191;					
RA	Chua K.Y., Doyle C.R., Simpson R.J., Turner K.J., Stewart G.A.,					
RA	Thomas W.R.;					
RT	"Isolation of cDNA coding for the major mite allergen Der p II by IgE					
RT	plaque immunoassay."					
RL	Int. Arch. Allergy Appl. Immunol. 91:118-123(1990).					
RN	[2]					
RP	SEQUENCE FROM N.A., AND VARIANTS					
RX	MEDLINE=21290932; PubMed=11398075;					
RA	Smith W.-A., Hales B.J., Jarnicki A.G., Thomas W.R.;					
RT	"Allergens of wild house dust mites: environmental Der p 1 and Der p 2					
RT	sequence polymorphisms."					
RL	J. Allergy Clin. Immunol. 107:985-992(2001).					
RN	[3]					
RP	PARTIAL SEQUENCE OF 18-57.					
RX	MEDLINE=89278484; PubMed=2732406;					
RA	Heymann P.W., Chapman M.D., Aalberse R.C., Fox J.W.,					
RA	Platts-Mills T.A.;					
RT	"Antigenic and structural analysis of group II allergens (Der f II					
RT	and Der p II) from house dust mites (Dermatophagoides spp).";					
RL	J. Allergy Clin. Immunol. 83:1055-1067(1989).					
RN	[4]					
RP	STRUCTURE BY NMR.					
RX	MEDLINE=98409423; PubMed=9737847;					
RA	Mueller G.A., Benjamin D.C., Rule G.S.;					
RT	"Tertiary structure of the major house dust mite allergen Der p 2:					
RT	sequential and structural homologies."					
RL	Biochemistry 37:12707-12714(1998).					
CC	-I- SUBCELLULAR LOCATION: Secreted.					
CC	-I- SIMILARITY: BELONGS TO THE NPC2 FAMILY.					
CC	-----					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC	the European Bioinformatics Institute. There are no restrictions on its					
CC	use by non-profit institutions as long as its content is in no way					
CC	modified and this statement is not removed. Usage by and for commercial					
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/					
CC	or send an email to license@isb-sib.ch).					
CC	-----					
DR	EMBL; AF276239; AAF86462.1; -.					



DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Tetanus toxin precursor (EC 3.4.24.68) (Tentoxylisin).  
OS Clostridium tetani.  
OG Plasmid.  
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
OC Clostridium.  
OX NCBI\_TaxID=1513;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87053814; PubMed=3536478;  
RA Eisel U., Jarausch W., Goretzki K., Henschen A., Engels J.,  
RA Weller U., Hudel M., Habermann E., Niemann H.;  
RT "Tetanus toxin: primary structure, expression in *E. coli*, and  
RT homology with botulinum toxins.";  
RL EMBO J. 5:2495-2502(1986).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX STRAIN=CN3911;  
RX MEDLINE=87040747; PubMed=3774547;  
RA Fairweather N.F., Lyness V.A.;  
RT "The complete nucleotide sequence of tetanus toxin.";  
RL Nucleic Acids Res. 14:7809-7812(1986).  
RN [3]  
RP SEQUENCE OF 742-1314 FROM N.A.  
RX MEDLINE=86085672; PubMed=3510187;  
RA Fairweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.;  
RT "Cloning, nucleotide sequencing, and expression of tetanus toxin  
RT fragment C in *Escherichia coli*.";  
RL J. Bacteriol. 165:21-27(1986).  
RN [4]  
RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.  
RX MEDLINE=90201034; PubMed=2108021;  
RA Krieglstein K., Henschen A., Weller U., Habermann E.;  
RT "Arrangement of disulfide bridges and positions of sulphydryl groups  
RT in tetanus toxin.";  
RL Eur. J. Biochem. 188:39-45(1990).  
RN [5]  
RP PARTIAL SEQUENCE.  
RX MEDLINE=92037649; PubMed=1935979;  
RA Krieglstein K.G., Henschen A.H., Weller U., Habermann E.;  
RT "Limited proteolysis of tetanus toxin. Relation to activity and  
RT identification of cleavage sites.";  
RL Eur. J. Biochem. 202:41-51(1991).  
RN [6]  
RP IDENTIFICATION AS ZINC-PROTEASE.  
RX MEDLINE=93010948; PubMed=1396558;  
RA Schiavo G., Poulain B., Rossetto O., Benfenati F., Tauc L.,  
RA Montecucco C.;  
RT "Tetanus toxin is a zinc protein and its inhibition of  
RT neurotransmitter release and protease activity depend on zinc.";  
RL EMBO J. 11:3577-3583(1992).  
RN [7]  
RP IDENTIFICATION OF SUBSTRATE.  
RX MEDLINE=93063293; PubMed=1331807;  
RA Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,  
RA Dasgupta B.R., Montecucco C.;  
RT "Tetanus and botulinum-B neurotoxins block neurotransmitter release  
RT by proteolytic cleavage of synaptobrevin.";  
RL Nature 359:832-835(1992).  
RN [8]  
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 874-1314.  
RX MEDLINE=97475217; PubMed=9334741;  
RA Umland T.C., Wingert L.M., Swaminathan S., Furey W.F., Schmidt J.J.,  
RA Sax M.;  
RT "Structure of the receptor binding fragment HC of tetanus  
RT neurotoxin.";  
RL Nat. Struct. Biol. 4:788-792(1997).  
CC -1- FUNCTION: TETANUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER  
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED  
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD  
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT  
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC

CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 76-GLN-1-PHE-77  
CC BOND OF SYNAPTOBREVIN-2  
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF 76-GLN-1-PHE-77 BOND IN  
CC SYNAPTOBREVIN.  
CC -1- SUBUNIT: THE PRECURSOR POLYPEPTIDE IS SUBSEQUENTLY CLEAVED TO  
CC YIELD SUBCHAINS L AND H. THESE REMAIN LINKED BY A DISULFIDE BRIDGE  
CC AND ARE NON-TOXIC AFTER SEPARATION.  
CC -1- MISCELLANEOUS: THE C-TERMINAL OF THE HEAVY CHAIN BINDS TO  
CC GANGLIOSIDE RECEPTORS.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X04436; CAA28033.1; -  
DR EMBL: M12739; AAA23282.1; -  
DR EMBL: X06214; CAA29564.1; -  
DR PIR: A25689; BTCLTN.  
DR PDB: 1AF9; 29-APR-98.  
DR PDB: 1A8D; 14-OCT-98.  
DR MEROPS: M27.001; -  
DR InterPro: IPR000395; Bontoxilysin.  
DR InterPro: IPR000130; Zn\_Mtpeptidse.  
DR Pfam: PF01742; Peptidase\_M27; 1.  
DR PRINTS: PR00760; BONTOXILYSIN.  
DR PRODOM: PD001963; Bontoxilysin; 1.  
DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc; Plasmid;  
KW 3D-structure.  
FT INIT\_MET 0 456  
FT CHAIN 1 456  
FT CHAIN 457 1314  
FT CHAIN 232 232  
FT METAL 233 233  
FT ACT\_SITE 233 233  
FT METAL 236 236  
FT TRANSMEM 226 246  
FT TRANSMEM 669 689  
FT DISULFID 438 466  
FT DISULFID 1076 1092  
SQ SEQUENCE 1314 AA; 150550 MW; 134C3657133EF81D CRC64;  
Query Match 40.9%; Score 74; DB 1; Length 1314;  
Best Local Similarity 100.0%; Pred.No.0.0083;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 QYIKANSKFIGITEL 15  
DB 829 QYIKANSKFIGITEL 843  
RESULT 4  
ALL2\_EURMA  
ID ALL2\_EURMA STANDARD; PRT; 145 AA.  
AC Q97ZZ2; O96430;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Mite group 2 allergen Eur m 2 precursor (Eur m 2.0101/Eur m 2.0102).  
GN EURN2.  
OS Euroglyphus maynei (House-dust mite).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
OC Acariformes; Sarcotiformes; Astigmata; Analgoidea; Pyroglyphidae;  
OC Euroglyphus.  
OX NCBI\_TaxID=6958;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99126275; PubMed=9925958;  
RA Smith W., Mills K., Hazell L., Hart B.J., Thomas W.;



RT "Molecular analysis of the group 1 and 2 allergens from the house dust  
RL mite, Euroglyphus maynei.", 118:15-22(1999).  
CC Int. Arch. Allergy Immunol. Secreted (By similarity).  
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).  
CC -!- SIMILARITY: BELONGS TO THE NPC2 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL; AF047613; AAC82349.1; -.  
CC EMBL; AF047614; AAC82350.1; -.  
CC HSSP; P49278; 1A9V.  
CC InterPro: IPR003172; EL\_DerP2\_DerF2.  
CC Pfam; PF02221; EL\_DerP2\_DerF2; 1.  
CC Allergen; Signal; Polymorphism.  
CC -----  
CC SIGNAL 1 16 POTENTIAL.  
CC CHAIN 17 145 MITE GROUP 2 ALLERGEN EUR M 2.  
CC DISULFID 24 135 BY SIMILARITY.  
CC DISULFID 37 43 BY SIMILARITY.  
CC DISULFID 89 94 BY SIMILARITY.  
CC VARIANT 21 21 I -> V (IN EUR M 2 0102).  
CC SEQUENCE 145 AA; 15747 MW; 6655B16C8503A565 CRC64;  
CC -----  
CC Query Match 38.1%; Score 69; DB 1; Length 145;  
CC Best Local Similarity 75.0%; Pred. No. 0.0055;  
CC Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
CC -----  
CC Y 17 GCHGSEPCNIHRGKPF 32  
CC II IIIII IIII I  
CC Db 36 GCKGSEPCVIHRGTAF 51  
CC -----  
CC RESULT 5  
CC PPNK\_HELPJ  
CC ID PPNK\_HELPJ STANDARD; PRT; 284 AA.  
CC AC 09ZJ81;  
CC DT 16-OCT-2001 (Rel. 40, Created)  
CC DT 16-OCT-2001 (Rel. 40, Last sequence update)  
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)  
CC DE Probable inorganic polyphosphate/ATP-NAD kinase (EC 2.7.1.23)  
CC DE Poly(P)/ATP NAD kinase).  
CC GN PPNK OR JHP1433.  
CC OS Helicobacter pylori J99 (Campylobacter pylori J99).  
CC OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
CC OC Helicobacter.  
CC OX NCBI\_TaxID=85963;  
CC RN [1]  
CC RP SEQUENCE FROM N.A.  
CC RX MEDLINE=99120557; PubMed=9923682;  
CC RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,  
CC RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carnel G.,  
CC RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,  
CC RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,  
CC RA Trust T.J.;  
CC RT "Genomic sequence comparison of two unrelated isolates of the human  
CC RT gastric pathogen Helicobacter pylori";  
CC RL Nature 397:176-180(1999).  
CC -----  
CC -!- FUNCTION: Catalyzes the phosphorylation of NAD to NADP. Utilizes  
CC ATP and other nucleoside triphosphates as well as inorganic  
CC polyphosphate as a source of phosphorus (By similarity).  
CC -!- CATALYTIC ACTIVITY: ATP + NAD(+) = ADP + NADP(+).  
CC -!- COFACTOR: Requires divalent metal ions for activity (By  
CC similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: BELONGS TO THE NAD KINASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between, the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL; AE001566; AAD07009.1; ALT\_INIT.  
CC DR InterPro: IPR002504; ATP\_NADK.  
CC DR Pfam; PF01513; NAD\_kinase; 1.  
CC KW Transferase; Kinase; NAD; NADP; Complete proteome.  
CC SQ SEQUENCE 284 AA; 31447 MW; E7E4056AC802BA7C CRC64;  
CC -----  
CC Query Match 28.2%; Score 51; DB 1; Length 284;  
CC Best Local Similarity 27.6%; Pred. No. 4.3;  
CC Matches 8; Conservative 8; Mismatches 13; Indels 0; Gaps 0;  
CC -----  
CC Y 1 OVIKANSKEIGITELGCGCHGSEPCNIHRG 29  
CC Db 140 EIVIAKKKALGVLDIKACAGHTPFNTYKG 168  
CC -----  
CC RESULT 6  
CC IBP6\_MOUSE  
CC ID IBP6\_MOUSE STANDARD; PRT; 238 AA.  
CC AC P47880;  
CC DT 01-FEB-1996 (Rel. 33, Created)  
CC DT 01-FEB-1996 (Rel. 33, Last sequence update)  
CC DT 15-DEC-1998 (Rel. 37, Last annotation update)  
CC DE Insulin-like growth factor binding protein 6 precursor (IGFBP-6)  
CC DE (IBP-6) (IGF-binding protein 6).  
CC GN IGFBP6 OR IGFBP-6.  
CC OS Mus musculus (Mouse).  
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
CC OX NCBI\_TaxID=10090;  
CC RN [1]  
CC RP SEQUENCE FROM N.A.  
CC RC TISSUE=Liver;  
CC RX MEDLINE=95121750; PubMed=7529732;  
CC RA Schuller A.G.P., Groffen C., van Neck J.W., Zwarthoff E.C.,  
CC RA Drop S.L.S.;  
CC RT "cDNA cloning and mRNA expression of the six mouse insulin-like  
CC RT growth factor binding proteins";  
CC RL Mol. Cell. Endocrinol. 104:57-66(1994).  
CC -!- FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFS  
CC AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH  
CC PROMOTING EFFECTS OF THE IGFS ON CELL CULTURE. THEY ALTER THE  
CC INTERACTION OF IGFS WITH THEIR CELL SURFACE RECEPTORS.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- PTM: O-GLYCOSYLATED.  
CC -!- SIMILARITY: CONTAINS 1 THYROGLOBULIN TYPE-I DOMAIN.  
CC -!- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING  
CC PROTEIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL; X81584; CAA57274.1; -.  
CC HSSP; P24593; IBOE.  
CC MGD; MGI:96441; Igfbp6.  
CC InterPro: IPR000867; Insl\_gro\_fac-pr.  
CC DR InterPro: IPR000716; Thyroglobulin\_1.  
CC DR Pfam; PF00086; thyroglobulin\_1.  
CC DR Pfam; PF00219; IGFBP; 1.  
CC DR SMART; SM00121; IB; 1.  
CC DR SMART; SM00211; TV; 1.  
CC DR PROSITE; PS00222; IGF\_BINDING; FALSE\_NEG.  
CC DR PROSITE; PS00484; THYROGLOBULIN\_1; 1.

KW Growth factor binding; signal; Glycoprotein.  
 FT SIGNAL 1 25 BY SIMILARITY.  
 FT CHAIN 26 238 INSULIN-LIKE GROWTH FACTOR BINDING  
 FT PROTEIN 6.  
 FT DOMAIN 184 232 THYROGLOBULIN TYPE I.  
 FT DISULFID 30 33 BY SIMILARITY.  
 FT DISULFID 41 45 BY SIMILARITY.  
 FT DISULFID 58 64 BY SIMILARITY.  
 FT DISULFID 72 85 BY SIMILARITY.  
 FT DISULFID 79 105 BY SIMILARITY.  
 FT DISULFID 160 188 BY SIMILARITY.  
 FT DISULFID 199 210 BY SIMILARITY.  
 FT DISULFID 212 232 BY SIMILARITY.  
 SQ SEQUENCE 238 AA; 25402 MW; CCEDCD1D6AD9D59F9 CRC64;  
 Query Match 27.1%; Score 49; DB 1; Length 238;  
 Best Local Similarity 39.3%; Pred. No. 7;  
 Matches 11; Conservative 3; Mismatches 12; Indels 2; Gaps 1;  
 QY 5 ANSKFIGITELGGC--HGSEPCNTHRGK 30  
 DB 51 AGSPADGCTEAGGCLRRGGQPCGVISPK 78  
 RESULT 7  
 LPXA\_CHRV1  
 ID LPXA\_CHRV1 STANDARD; PRT; 259 AA.  
 AC Q46481;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Acyl-[acyl-carrier-protein]-UDP-N-acetylglucosamine O-acyltransferase  
 DE (EC 2.3.1.129) (UDP-N-acetylglucosamine acyltransferase).  
 GN LPXA.  
 OS Chromatium vinosum.  
 OC Bacteria; Proteobacteria; gamma subdivision; Chromatiaceae;  
 OC Allochrochromium.  
 OX NCBI\_TaxID=1049;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D / ATCC 17899 / DSM 180;  
 RA Chen Y.L., Knaff D.B.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: INVOLVED IN THE BIOSYNTHESIS OF LIPID A, A  
 CC PHOSPHORYLATED GLYCOPOLIPID THAT ANCHORS THE LIPOPOLYSACCHARIDE TO  
 CC THE OUTER MEMBRANE OF THE CELL (BY SIMILARITY).  
 CC -!- CATALYTIC ACTIVITY: (R)-3-hydroxytetradecanoyl-[acyl-carrier  
 CC protein] + UDP-N-acetylglucosamine = [acyl-carrier protein] +  
 CC UDP-3-O-(3-hydroxytetradecanoyl)-N-acetylglucosamine.  
 CC -!- PATHWAY: Lipid A biosynthesis; first step.  
 CC -!- SUBUNIT: HOMOTRIMER (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: BELONGS TO THE TRANSFERASE HEXAPEPTIDE REPEAT FAMILY.  
 CC LPXA SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L76417; AAB02979.1; -;  
 DR HSP; P10440; ILXA.  
 DR InterPro; IPR001451; Hexapep.transf.  
 DR Pfam; PF00132; hexapep; 5.  
 DR PROSITE; PS00101; HEXAPEP\_TRANSFERASES; FALSE\_NEG.  
 KW Transferase; Acyltransferase; Lipid A biosynthesis; Lipid synthesis;  
 KW Repeat.  
 SQ SEQUENCE 259 AA; 28172 MW; D7F4690066180CDA CRC64;  
 Query Match 27.1%; Score 49; DB 1; Length 259;  
 Best Local Similarity 37.9%; Pred. No. 7.6;  
 Matches 11; Conservative 5; Mismatches 8; Gaps 2;  
 QY 3 IRANSKFITGTELGG----CH----GSEP 23  
 DB 39 IESNVRFICVTRMGAAHNRVCHGATLGLSEP 67  
 RESULT 8  
 URH1\_YEAST  
 ID URH1\_YEAST STANDARD; PRT; 378 AA.  
 AC Q04179;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Uridine nucleosidase (EC 3.2.2.3) (Uridine ribohydrolase).  
 GN URH1 OR YDR400W OR D9509.19.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.  
 RX MEDLINE=99431800; PubMed=10501935;  
 RA Kurtz J.-E., Exinger F., Erbs P., Jund R.;  
 RT "New insights into the pyrimidine salvage pathway of Saccharomycetes  
 RT cerevisiae: requirement of six genes for cytidine metabolism.";  
 RL Curr. Genet. 36:130-136(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E.,  
 RA Berno A., Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,  
 RA Hunkle-Smith S., Hyman R., Komp C., Lashkari D., Lew H., Lin D.,  
 RA Mosedale D., Nakahara K., Namath A., Oefner P., Oh C., Petel F.X.,  
 RA Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N.,  
 RA Winant A., Yelton M., Botstein D., Davis R.W.;  
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: ALSO ACTS ON CYTIDINE.  
 CC -!- CATALYTIC ACTIVITY: Uridine + H(2)O = uracil + D-ribose.  
 CC -!- SIMILARITY: BELONGS TO THE IUNH FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF217406; AAG44107.1; -;  
 DR EMBL; U32274; AAB64841.1; -;  
 DR HSP; Q27546; 1MAS.  
 DR SGD; S0002808; URH1.  
 DR InterPro; IPR001910; I/U\_NHdase.  
 DR Pfam; PF01156; IU\_nuc\_hydro; 1.  
 DR ProDom; PD007736; I/U\_NHdase; 1.  
 DR PROSITE; PS01247; IUNH; 1.  
 KW Hydrolase; Glycosidase.  
 FT ACT\_SITE 292 292 BY SIMILARITY.  
 SQ SEQUENCE 378 AA; 42317 MW; 58499E2B7480BEE5 CRC64;  
 Query Match 26.8%; Score 48.5; DB 1; Length 378;  
 Best Local Similarity 45.8%; Pred. No. 13;  
 Matches 11; Conservative 3; Mismatches 9; Indels 1; Gaps 1;  
 QY 2 YIKANSKFITGTELGGCHGSEPCN 25  
 DB 183 YLKKSVKYISMG-GGLHGLGNCN 205  
 RESULT 9  
 GIDA\_MYCGE  
 ID GIDA\_MYCGE STANDARD; PRT; 612 AA.

```

AC P47619;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glucose inhibited division protein A.
GN GIDA OR MG379.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Ufferback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium.";
RL Science 270:397-403(1995).
RN [2]
RP SEQUENCE OF 301-394 FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=94075230; PubMed=8253680;
RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;
RT "A survey of the Mycoplasma genitalium genome by using random
sequencing.";
RL J. Bacteriol. 175:7918-7930(1993).
CC -1- FUNCTION: NOT KNOWN.
CC -1- SIMILARITY: BELONGS TO THE GIDA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U39719; AAC71606.1; -
CC EMBL; U01812; AAD12347.1; -
CC TIGR; MG379; -
CC InterPro; IPR002218; GIDA.
CC InterPro; IPR004416; GIDA_sub.
CC InterPro; IPR000205; NAD_binding.
CC Pfam; PF01134; GIDA_1.
CC ProDom; PD003738; GIDA; 1.
CC TIGRFAMS; TIGR00136; gida; 1.
CC PROSITE; PS01280; GIDA_1; 1.
CC PROSITE; PS01281; GIDA_2; 1.
CC Complete proteome.
KW SEQUENCE 612 AA; 68995 MW; 67A3F64C985B5912 CRC64;
Query Match 26.5%; Score 48; DB 1; Length 612;
Best Local Similarity 41.4%; Pred. No. 24;
Matches 12; Conservative 1; Mismatches 12; Indels 4; Gaps 1;

QY 3 IKANSKFIGITELGGCHSEPCNTHRGKP 31
| | | : | | | | | | | |
Db 26 IKVNLVLDINHLCG-----PCNPSIGGP 50

RESULT 10
Y008_UREPA
ID Y008_UREPA STANDARD; PRT; 158 AA.
AC Q9PRD7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein U0008.
GN U0008.
OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).

```

```

OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.
OX NCBI_TaxID=134821;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Serovar 3;
RX MEDLINE=20500219; PubMed=11048724;
RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,
RA Cassell G.H.;
RT "The complete sequence of the mucosal pathogen Ureaplasma
urealyticum.";
RL Nature 407:757-762(2000).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: TO U.PARVUM U0007, U0041 AND U0042.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE002100; AAF30413.1; -
CC KW Hypothetical protein; Transmembrane; Complete proteome.
CC TRANSMEM 45 65 POTENTIAL.
CC TRANSMEM 76 96 POTENTIAL.
CC TRANSMEM 106 126 POTENTIAL.
CC SEQUENCE 158 AA; 16990 MW; ACP974AA54936969 CRC64;
Query Match 26.0%; Score 47; DB 1; Length 158;
Best Local Similarity 45.8%; Pred. No. 9.3;
Matches 11; Conservative 3; Mismatches 4; Indels 6; Gaps 1;

QY 4 KANSKFIGITELGGCHSEPCNTH 27
| | | : | | | | | | | |
Db 131 KAKAAGMTEL-----EYCNLH 148

RESULT 11
RFE_MYCLE
ID RFE_MYCLE STANDARD; PRT; 398 AA.
AC P45830;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative undecaprenyl-phosphate alpha-N-acetylglucosaminyltransferase
DE (EC 2.4.1.-).
GN RFE OR ML1137.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RA Smith D.R., Robison K.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies K.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
CC -1- CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + undecaprenyl
monophosphate = UMP + undecaprenyl N-acetyl-alpha-D-glucosaminyl

```

CC CC pyrophosphate.  
 CC -1- COFACTOR: Magnesium and manganese (By similarity).  
 CC -1- PATHWAY: Lipopolysaccharide biosynthesis.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 4. WECA  
 CC SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U15186; AAA63094.1; -;  
 CC EMBL; AL583920; CAC31518.1; ALT\_INIT.  
 CC Leproma; ML1137; -;  
 CC InterPro: IPR00715; Glycos\_transf\_4.  
 CC Pfam; PF00953; Glycos\_transf\_4; 1.  
 CC Lipopolysaccharide biosynthesis; Glycosyltransferase; Transferase;  
 CC Transmembrane; Complete proteome.  
 CC TRANSMEM 33 53 POTENTIAL.  
 CC TRANSMEM 79 99 POTENTIAL.  
 CC TRANSMEM 105 125 POTENTIAL.  
 CC TRANSMEM 148 168 POTENTIAL.  
 CC TRANSMEM 175 195 POTENTIAL.  
 CC TRANSMEM 198 218 POTENTIAL.  
 CC TRANSMEM 225 245 POTENTIAL.  
 CC TRANSMEM 260 280 POTENTIAL.  
 CC TRANSMEM 295 315 POTENTIAL.  
 CC TRANSMEM 347 367 POTENTIAL.  
 CC TRANSMEM 372 392 POTENTIAL.  
 CC SEQUENCE 398 AA; 42088 MW; 37564293606CF9A6 CRC64;  
 CC  
 CC Query Match 26.0%; Score 47; DB 1; Length 398;  
 CC Best Local Similarity 43.5%; Pred. No. 22;  
 CC Matches 10; Conservative 1; Mismatches 12; Indels 0; Gaps 0;  
 CC  
 CC QY 10 IGITELGCGHSEPCNIHRGKPF 32  
 CC : : | | | | | | | |  
 CC Db 233 ISVVLGAGSLGFLPHNFRKIF 255  
 CC : : | | | | | | | |  
 CC  
 CC RESULT 12  
 CC RFE\_MYCTU  
 CC ID RFE\_MYCTU STANDARD; PRT; 404 AA.  
 CC AC Q10606;  
 CC DT 01-OCT-1996 (Rel. 34, Created)  
 CC DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 CC DE Putative undecaprenyl-phosphate alpha-N-acetylglucosaminyltransferase  
 CC DE (EC 2.4.1.-).  
 CC GN RFE OR RV1302 OR MT1341 OR MTCY373.22.  
 CC OS Mycobacterium tuberculosis.  
 CC OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;  
 CC OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 CC OX NCBI\_TaxID=1773;  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC RP STRAIN=H37RV;  
 CC RX MEDLNP=98295987; PubMed=9634230;  
 CC RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 CC Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,  
 CC Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 CC Davies K., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
 CC Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 CC Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 CC Rutter S., Seeger K., Skellon S., Squares S., Squares R.,  
 CC Sulston J.E., Taylor K., Whitehead S., Barrall B.G.;  
 CC RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 CC complete genome sequence."  
 CC RL Nature 393:537-544(1998).

CC  
 CC [2]  
 CC SEQUENCE FROM N.A.  
 CC RP STRAIN=CDC 1551 / Oshkosh;  
 CC RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 CC Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
 CC Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
 CC Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
 CC Bishai W.;  
 CC "Whole genome comparison of Mycobacterium tuberculosis clinical and  
 CC laboratory strains."  
 CC RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + undecaprenyl  
 CC monophosphate = UMP + undecaprenyl N-acetyl-alpha-D-glucosaminyl  
 CC pyrophosphate.  
 CC -1- COFACTOR: Magnesium and manganese (By similarity).  
 CC -1- PATHWAY: Lipopolysaccharide biosynthesis.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 4. WECA  
 CC SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; Z73419; CAA97735.1; -;  
 CC EMBL; AF007008; AAK45603.1; ALT\_INIT.  
 CC TIGR; MT1341; -;  
 CC TubercuList; RV1302; -;  
 CC InterPro: IPR000715; Glycos\_transf\_4.  
 CC Pfam; PF00953; Glycos\_transf\_4; 1.  
 CC KW Lipopolysaccharide biosynthesis; Glycosyltransferase; Transferase;  
 CC Transmembrane; Complete proteome.  
 CC TRANSMEM 33 53 POTENTIAL.  
 CC TRANSMEM 79 99 POTENTIAL.  
 CC TRANSMEM 105 125 POTENTIAL.  
 CC TRANSMEM 148 168 POTENTIAL.  
 CC TRANSMEM 175 195 POTENTIAL.  
 CC TRANSMEM 198 218 POTENTIAL.  
 CC TRANSMEM 225 245 POTENTIAL.  
 CC TRANSMEM 259 279 POTENTIAL.  
 CC TRANSMEM 295 315 POTENTIAL.  
 CC TRANSMEM 347 367 POTENTIAL.  
 CC TRANSMEM 372 392 POTENTIAL.  
 CC SEQUENCE 404 AA; 42257 MW; 57D7D2807034A426 CRC64;  
 CC  
 CC Query Match 26.0%; Score 47; DB 1; Length 404;  
 CC Best Local Similarity 43.5%; Pred. No. 23;  
 CC Matches 10; Conservative 1; Mismatches 12; Indels 0; Gaps 0;  
 CC  
 CC QY 10 IGITELGCGHSEPCNIHRGKPF 32  
 CC : : | | | | | | | |  
 CC Db 233 ISVVLGAGSLGFLPHNFRKIF 255  
 CC : : | | | | | | | |  
 CC  
 CC RESULT 13  
 CC BHMT\_HUMAN  
 CC ID BHMT\_HUMAN STANDARD; PRT; 406 AA.  
 CC AC Q93088;  
 CC DT 01-NOV-1997 (Rel. 35, Created)  
 CC DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 CC DE Betaine--homocysteine S-methyltransferase (EC 2.1.1.5).  
 CC GN BHMT.  
 CC OS Homo sapiens (Human).  
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 CC OX NCBI\_TaxID=9606;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.

```
RC TISSUE=Liver;
RX MEDLINE=96394355; PubMed=8798461;
RA Garrow T.A.;
RT "Purification, kinetic properties, and cDNA cloning of mammalian
  betaine-homocysteine methyltransferase.";
RL J. Biol. Chem. 271:22831-22838(1996).
RN [2]
RP TISSUE SPECIFICITY.
RX MEDLINE=97428198; PubMed=9281325;
RA Sundén S.L.F., Renduchintala M.S., Park E.I., Miklasz S.D.,
  Garrow T.A.;
RT "Betaine-homocysteine methyltransferase expression in porcine and
  human tissues and chromosomal localization of the human gene.";
RL Arch. Biochem. Biophys. 345:171-174(1997).
RN [3]
RP CHARACTERIZATION.
RC TISSUE=Liver;
RX MEDLINE=98348402; PubMed=9681996;
RA Millian N.S., Garrow T.A.;
RT "Human betaine-homocysteine methyltransferase is a zinc
  metalloenzyme.";
RL Arch. Biochem. Biophys. 356:93-98(1998).
CC -!- FUNCTION: INVOLVED IN THE REGULATION OF HOMOCYSTEINE METABOLISM.
  CONVERTS BETAIN AND HOMOCYSTEINE TO DIMETHYLGLYCINE AND
  METHIONINE, RESPECTIVELY. THIS REACTION IS ALSO REQUIRED FOR THE
  IRREVERSIBLE OXIDATION OF CHOLINE.
CC -!- CATALYTIC ACTIVITY: Trimethylammonioacetate + L-homocysteine =
  dimethylglycine + L-methionine.
CC -!- COFACTOR: ZINC.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: FOUND EXCLUSIVELY IN LIVER AND KIDNEY.
CC -!- DISEASE: DEFECTS IN BHMT COULD LEAD TO HYPERHOMOCYSTEINEMIA. BUT
  SUCH A DEFECT HAS NOT YET BEEN OBSERVED. HYPERHOMOCYSTEINEMIA IS
  AN INDEPENDENT RISK FACTOR FOR THE DEVELOPMENT OF ARTERIOSCLEROTIC
  VASCULAR DISEASE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U50929; AAC50668.1; .
DR Genew; HGNC:1047; BHMT.
DR MIM; 602888; .
DR InterPro; IPR003726; S_methyl_trans.
DR Pfam; PF02574; S-methyl_trans; 1.
DR Transferase; Methyltransferase; Zinc.
DR KW SEQUENCE 406 AA; 44970 MW; 79492DE7F1DEFE9A CRC64;
SQ
  Query Match 26.0%; Score 47; DB 1; Length 406;
  Best Local Similarity 47.1%; Pred. No. 23;
  Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
  YQ 10 IGITELGGCHGSEPCNI 26
  : : :||| | | |
  Db 291 LGVRYIGCCGFEFYHI 307
  RESULT 14
  BHMT_MOUSE
  ID BHMT_MOUSE STANDARD; PRT; 407 AA.
  AC O35490;
  DT 15-JUL-1999 (Rel. 38, Created)
  DT 15-JUL-1999 (Rel. 38, Last sequence update)
  DT 30-MAY-2000 (Rel. 39, Last annotation update)
  DE Betaine--homocysteine S-methyltransferase (EC 2.1.1.5).
  GN BHMT.
  OS Mus musculus (Mouse).
  OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
  CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
  CC
```

```
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Sowden M.P., Smith H.C.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INVOLVED IN THE REGULATION OF HOMOCYSTEINE METABOLISM.
  CONVERTS BETAIN AND HOMOCYSTEINE TO DIMETHYLGLYCINE AND
  METHIONINE, RESPECTIVELY. THIS REACTION IS ALSO REQUIRED FOR THE
  IRREVERSIBLE OXIDATION OF CHOLINE.
CC -!- CATALYTIC ACTIVITY: Trimethylammonioacetate + L-homocysteine =
  dimethylglycine + L-methionine.
CC -!- COFACTOR: ZINC (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF033381; AAB87501.1; .
DR MGD; MGI:1339972; Bhmt.
DR InterPro; IPR003726; S_methyl_trans.
DR Pfam; PF02574; S-methyl_trans; 1.
DR KW Transferase; Methyltransferase; Zinc.
DR SQ SEQUENCE 407 AA; 45020 MW; 1C600B99CC44EE32 CRC64;
  Query Match 26.0%; Score 47; DB 1; Length 407;
  Best Local Similarity 47.1%; Pred. No. 23;
  Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
  YQ 10 IGITELGGCHGSEPCNI 26
  : : :||| | | |
  Db 291 LGVRYIGCCGFEFYHI 307
  RESULT 15
  BHMT_RAT
  ID BHMT_RAT STANDARD; PRT; 407 AA.
  AC O09171;
  DT 15-DEC-1998 (Rel. 37, Created)
  DT 15-DEC-1998 (Rel. 37, Last sequence update)
  DT 15-JUN-2002 (Rel. 41, Last annotation update)
  DE Betaine--homocysteine S-methyltransferase (EC 2.1.1.5).
  GN BHMT.
  OS Rattus norvegicus (Rat).
  OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
  CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
  CC NCBI_TaxID=10116;
  RN [1]
  RP SEQUENCE FROM N.A.
  RA STRAIN=Sprague-Dawley; TISSUE=Liver;
  RX MEDLINE=96354796; PubMed=8753772;
  RA Forestier M., Reichen J., Solioz M.;
  RT "Application of mRNA differential display to liver cirrhosis: reduced
  fetuin expression in biliary cirrhosis in the rat.";
  RL Biochem. Biophys. Res. Commun. 225:377-383(1996).
  RN [2]
  RP SEQUENCE FROM N.A.
  RX MEDLINE=99348027; PubMed=10417327;
  RA Sowden M.P., Collins H.L., Smith H.C., Garrow T.A., Sparks J.D.,
  RA Sparks C.E.;
  RT "Apolipoprotein B mRNA and lipoprotein secretion are increased in
  McArdle RH-7777 cells by expression of betaine-homocysteine
  S-methyltransferase.";
  RL Biochem. J. 341:639-645(1999).
  CC -!- FUNCTION: INVOLVED IN THE REGULATION OF HOMOCYSTEINE METABOLISM.
  CONVERTS BETAIN AND HOMOCYSTEINE TO DIMETHYLGLYCINE AND
  METHIONINE, RESPECTIVELY. THIS REACTION IS ALSO REQUIRED FOR THE
  IRREVERSIBLE OXIDATION OF CHOLINE.
```

```
CC -!- CATALYTIC ACTIVITY: Trimethylammonioacetate + L-homocysteine =
CC dimethylglycine + L-methionine.
CC -!- COFACTOR: ZINC (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U96133; AAB53763.1; -
CC EMBL: AF038870; AAB95481.1; -
CC InterPro: IPR003726; S_methyl_trans.
CC Pfam: PF02574; S-methyl_trans; 1.
CC KW Transferase; Methyltransferase; Zinc.
CC SQ SEQUENCE 407 AA; 44976 MW; 36E1D04ABE425887 CRC64;
CC -----
DR EMBL: U96133; AAB53763.1; -
DR EMBL: AF038870; AAB95481.1; -
DR InterPro: IPR003726; S_methyl_trans.
DR Pfam: PF02574; S-methyl_trans; 1.
KW KW Transferase; Methyltransferase; Zinc.
SQ SEQUENCE 407 AA; 44976 MW; 36E1D04ABE425887 CRC64;

Query Match 26.08; Score 47; DB 1; Length 407;
Best Local Similarity 47.18; Pred. No. 23;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 10 IGITELGGCHGSEPCNI 26
Db :|:|||||:|
291 LGVRYIGCCGFEPYHI 307

Search completed: December 4, 2002, 12:53:31
Job time : 4.46667 secs
```

GenCore version 5.1.3

Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 12:49:06 ; Search time 12.4 Seconds  
(without alignments)  
531.734 Million cell updates/sec

Title: US-09-362-731A-5

Perfect score: 181

Sequence: 1 QYKANSKFGITELGGCHGSEPCNIHRGKPF 32

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phase:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_rvirus:\*

16: sp\_bacteriap:\*

17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	91	50.3	170	5	Q9BIX2 dermatophag
2	89	49.2	129	5	Q8WQK5 dermatophag
3	74	40.9	1310	2	Q93N27 clostridium
4	57.5	31.8	458	2	Q9AKT9 paenibacill
5	53.5	29.6	410	10	O80919 arabidopsis
6	52	28.7	289	5	Q07338 pyrenotes tr
7	52	28.7	357	8	Q9B8W9 taenia cras
8	52	28.7	1159	16	Q8RCZ1 thermoanaer
9	52	28.7	1884	11	P70207 mus musculu
10	51	28.2	149	7	Q31495 oncorhynch
11	51	28.2	202	5	Q8SW20 encephalito
12	51	28.2	355	8	Q9TE80 echinococcu
13	51	28.2	355	8	Q9S8U8 echinococcu
14	51	28.2	355	8	Q953N0 echinococcu
15	50.5	27.9	60	7	Q31585 salmo salar
16	50.5	27.9	71	7	Q9XRJ9 salvelinus

17	50.5	27.9	85	7	Q95IS2
18	50.5	27.9	85	7	Q95HY1
19	50.5	27.9	86	7	Q95HX4
20	50.5	27.9	244	7	Q31590
21	50	27.6	358	2	Q93S04
22	50	27.6	372	8	Q9B8V1
23	50	27.6	1067	13	Q9YHU6
24	50	27.6	1328	4	Q9U1W1
25	50	27.6	1963	4	Q75051
26	50	27.6	3680	5	Q9VR08
27	49.5	27.3	67	7	Q31578
28	49.5	27.3	253	12	Q9YMU0
29	49.5	27.3	329	16	Q8XBX8
30	49	27.1	238	11	Q91X24
31	49	27.1	242	17	Q8TMH3
32	49	27.1	510	10	Q9LE20
33	48.5	26.8	202	8	Q9TIP4
34	48.5	26.8	204	8	O47109
35	48	26.5	62	2	Q9X5H5
36	48	26.5	134	17	Q8TIK8
37	48	26.5	363	11	Q9EQE8
38	48	26.5	363	11	Q91WS4
39	48	26.5	484	17	O58041
40	48	26.5	484	17	Q9UY44
41	48	26.5	502	10	O82752
42	48	26.5	505	10	Q9FJ87
43	48	26.5	532	10	Q9SLV0
44	48	26.5	817	5	Q21854
45	47.5	26.2	84	13	Q9DEK4

## ALIGNMENTS

RESULT 1

Q9BIX2 ID Q9BIX2 PRELIMINARY; PRT; 170 AA.

AC Q9BIX2;  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE Allergen Def f II (Fragment).  
 OS Dermatophagoides farinae (House-dust mite).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Acariformes; Sarcophagales; Astigmata; Analgoidea; Pyroglyphidae;  
 OC Dermatophagoides;  
 OX NCBI\_TaxID=6954;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Hao M., Xu J., Zhong N.;  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF346905; AAK30133.1; -  
 DR HSSP; Q00855; 1AHK.  
 DR InterPro; IPR003172; El\_DerF2\_DerF2.  
 DR Pfam; PF02221; El\_DerF2\_DerF2; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 170 AA; 18781 MW; 0C2B58734C9D443A CRC64;

Query Match 50.3%; Score 91; DB 5; Length 170;  
 Best Local Similarity 54.8%; Pred. No. 5.7e-06;  
 Matches 17; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 2 YIKANSKFGITELGGCHGSEPCNIHRGKPF 32

|| : | : ||||| : |||||

Db 46 YTHIANNEIKKVMVDGCHGSDPCIIHRGKPF 76

RESULT 2

Q8WQK5 ID Q8WQK5 PRELIMINARY; PRT; 129 AA.

AC Q8WQK5;

DT 01-MAR-2002 (Tremblrel. 20, Created)

DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)

DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 OS Major Der f 2 isoform (Fragment).  
 OS Dermatophagoides farinae (House-dust mite).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Acariformes; Sarcoptriformes; Astigmata; Analgoidea; Pyroglyphidae;  
 OC Dermatophagoides.  
 OX NCBI\_TaxID=6954;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Jin H.S., Oh S.H., Hong C.-S.;  
 RT "cDNA sequence encoding major Der f 2 isoform in Korea.";  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY066008; AAL47677.1; -;  
 DR InterPro; IPR003172; EL\_Derp2\_DerF2.  
 DR Pfam; PF02221; EL\_Derp2\_DerF2; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 129 AA; 14035 MW; 832F72E25FE4F43F CRC64;

Query Match 49.2%; Score 89; DB 5; Length 129;  
 Best Local Similarity 87.5%; Pred. No. 8.6e-06;  
 Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 GCHGSEPCNIHRGKPF 32  
 DB 20 GCHGSDPCIHRGKPF 35

RESULT 3  
 Q93N27

ID Q93N27 PRELIMINARY; PRT; 1310 AA.  
 AC Q93N27;  
 DT 01-DEC-2001 (TReMBLrel. 19, Created)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
 DE Tetanus toxin (Fragment).  
 OS Clostridium tetani.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;  
 OC Clostridiales; Clostridiaceae; Clostridium.  
 OX NCBI\_TaxID=1513;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Shumin Z., Dianliang L.;  
 RT "Cloning and sequence analysis of tetanus toxin gene.";  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF389424; AAK72964.2; -;  
 DR InterPro; IPR000395; Bontoxilysin.  
 DR InterPro; IPR001064; Crystallin.  
 DR InterPro; IPR000130; Zn\_Mtpeptidse.  
 DR Pfam; PF01742; Peptidase\_M27; 1.  
 DR ProDom; PD001963; Bontoxilysin; 1.  
 DR PROSITE; PS00225; CRYSTALLIN\_BETAGAMMA; UNKNOWN\_1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; UNKNOWN\_1.  
 FT NON\_TER 1  
 FT NON\_TER 1310 1310  
 SQ SEQUENCE 1310 AA; 150316 MW; 9EADDC914418E450 CRC64;

Query Match 40.9%; Score 74; DB 2; Length 1310;  
 Best Local Similarity 100.0%; Pred. No. 0.02;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15  
 DB 831 QYIKANSKFIGITEL 845

RESULT 4

Q9AKT9  
 ID Q9AKT9 PRELIMINARY; PRT; 458 AA.  
 AC Q9AKT9;  
 DT 01-JUN-2001 (TReMBLrel. 17, Created)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE Nifb protein.

GN NIFB.  
 OS Paenibacillus azotofixans.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
 OC Paenibacillaceae; Paenibacillus.  
 OX NCBI\_TaxID=44251;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 35681;  
 RA Choo Q.C., Othman A.S., Samian M.R., Najimudin N.;  
 RT "Cloning, DNA sequencing and phylogeny of two nifH-homologous genes  
 from Paenibacillus azotofixans.";  
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ299453; CAC27790.1; -;  
 DR InterPro; IPR003731; DUF153.  
 DR InterPro; IPR000385; MoaA\_NifB\_PqqE.  
 DR Pfam; PF02579; DUF153; 1.  
 DR Pfam; PF01444; MoaA\_NifB\_PqqE; 1.  
 DR TIGRFAMS; TIGR01290; nifB; 1.  
 DR PROSITE; PS01305; MOA\_NIFB\_PQOE; 1.  
 SQ SEQUENCE 458 AA; 50530 MW; 5A059C5414796CF3 CRC64;

Query Match 31.8%; Score 57.5; DB 2; Length 458;  
 Best Local Similarity 44.4%; Pred. No. 2.2;  
 Matches 12; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

QY 5 ANSKFTGITELGG-CHGSEPCNIHRGK 30  
 DB 368 AEVKGFIGIRKIQAYCHGTADNCGRGE 394

RESULT 5  
 O80919

ID O80919 PRELIMINARY; PRT; 410 AA.  
 AC O80919; O94K45;  
 DT 01-NOV-1998 (TReMBLrel. 08, Created)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE Expressed protein (Hypothetical 47.7 kDa protein).  
 OS AT2G38320.  
 GN Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Rounsley S.D., Lin X., Ketchum K.A., Crosby M.L., Brandon R.C.,  
 RA Sykes S.M., Kaul S., Mason T.M., Kerlavage A.R., Adams M.D.,  
 RA Somerville C.R., Venter J.C.;  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Lin X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Town C.D., Kaul S.;  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,  
 RA Goldsmith A.D., Lee J.M., Quach H.L., Tang C.C., Toriumi M., Yu G.,  
 RA Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J.,  
 RA Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koeseema E.,  
 RA Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M.,  
 RA Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,  
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;  
 RT "Full Length cDNA of gene TIGR21.19/At2g38320 (GI:3395440).";  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC004683; AAC28772.2; -;



[illegible]



```

OX NCBI_TaxID=6211;
RN [1]
RP SEQUENCE FROM N.A.
RA Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;
RT "Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases."
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
DR EMBL; AB018440; BAA84928.2; -.
DR InterPro; IPR000179; Cyt_b_b6.
DR Pfam; PF00032; cytochrome_b_c_1.
DR Pfam; PF00033; cytochrome_b_n; 1.
DR PROSITE; PS00193; CYTOCHROME_B_QQ; UNKNOWN_1.
DR Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;
KW Electron transport; Heme; Mitochondrion; Respiratory chain;
KW Transmembrane.
SQ SEQUENCE 355 AA; 41268 MW; C51D1A0EC7C6E948 CRC64;

Query Match 28.2%; Score 51; DB 8; Length 355;
Best Local Similarity 52.9%; Pred. No. 17;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 7 SKFIGITELGGCHGSEP 23
DB 317 SLFLSLTYLGGCHPEYP 333
| | : | | | | | |
| | : | | | | | |

RESULT 13
OY58U8 PRELIMINARY; PRT; 355 AA.
AC Q958U8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Cytochrome b.
GN CYTB.
OS Echinococcus granulosus.
OG Mitochondrion.
OC Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;
OC Cyclophyllidae; Taeniidae; Echinococcus.
OX NCBI_TaxID=6210;
RN [1]
RP SEQUENCE FROM N.A.
RA Le T.H., Blair D., Pearson M., McManus D.P.;
RT "Echinococcus mitochondrial genome."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
DR EMBL; AF297617; AAK51678.1; -.
DR InterPro; IPR001751; CaBP.s100.
DR Pfam; PF00032; cytochrome_b_c_1.
DR Pfam; PF00033; cytochrome_b_n; 1.
DR PROSITE; PS00193; CYTOCHROME_B_QQ; UNKNOWN_1.
DR PROSITE; PS00303; S100_CaBP; UNKNOWN_1.
DR Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;
KW Electron transport; Heme; Mitochondrion; Respiratory chain;
KW Transmembrane.
SQ SEQUENCE 355 AA; 41252 MW; D44ACFD58CAE84B7 CRC64;

Query Match 28.2%; Score 51; DB 8; Length 355;
Best Local Similarity 52.9%; Pred. No. 17;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 7 SKFIGITELGGCHGSEP 23
DB 317 SLFLSLTYLGGCHPEYP 333
| | : | | | | | |
| | : | | | | | |

RESULT 14
OY53N0 PRELIMINARY; PRT; 355 AA.
AC Q953N0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Cytochrome b.
GN CYTB.
OS Echinococcus granulosus.
OG Mitochondrion.
OC Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;
OC Cyclophyllidae; Taeniidae; Echinococcus.
OX NCBI_TaxID=6210;
RN [1]
RP SEQUENCE FROM N.A.
RA Le T.H., Dai T.H.N., Blair D., McManus D.P.;
RT "Complete mitochondrial genome of Echinococcus granulosus (G4 genotype)."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
DR EMBL; AF346403; AAK82343.1; -.
DR InterPro; IPR000179; Cyt_b_b6.
DR Pfam; PF00032; cytochrome_b_c_1.
DR Pfam; PF00033; cytochrome_b_n; 1.
DR PROSITE; PS00193; CYTOCHROME_B_QQ; UNKNOWN_1.
DR Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;
KW Electron transport; Heme; Mitochondrion; Respiratory chain;
KW Transmembrane.
SQ SEQUENCE 355 AA; 41220 MW; 61D849BF10DAED00 CRC64;

Query Match 28.2%; Score 51; DB 8; Length 355;
Best Local Similarity 52.9%; Pred. No. 17;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 7 SKFIGITELGGCHGSEP 23
DB 317 SLFLSLTYLGGCHPEYP 333
| | : | | | | | |
| | : | | | | | |

RESULT 15
OY1585 PRELIMINARY; PRT; 60 AA.
AC Q31585;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE (DB02) MHC class II beta 1 (Fragment).
OS Salmo salar (Atlantic salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX NCBI_TaxID=8030;
RN [1]
RP SEQUENCE FROM N.A.
RA Grimholt U., Olsaker I., de Vries Lindstrom C., Lie O.;
RT "A study of polymorphism in the MHC class II beta 1 and MHC class I
alpha 2 domain exons of Atlantic salmon (Salmo salar).";
```

```

OX NCBI_TaxID=6211;
RN [1]
RP SEQUENCE FROM N.A.
RA Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;
RT "Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases."
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
DR EMBL; AB018440; BAA84928.2; -.
DR InterPro; IPR000179; Cyt_b_b6.
DR Pfam; PF00032; cytochrome_b_c_1.
DR Pfam; PF00033; cytochrome_b_n; 1.
DR PROSITE; PS00193; CYTOCHROME_B_QQ; UNKNOWN_1.
DR Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;
KW Electron transport; Heme; Mitochondrion; Respiratory chain;
KW Transmembrane.
SQ SEQUENCE 355 AA; 41268 MW; C51D1A0EC7C6E948 CRC64;

Query Match 28.2%; Score 51; DB 8; Length 355;
Best Local Similarity 52.9%; Pred. No. 17;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 7 SKFIGITELGGCHGSEP 23
DB 317 SLFLSLTYLGGCHPEYP 333
| | : | | | | | |
| | : | | | | | |

RESULT 13
OY58U8 PRELIMINARY; PRT; 355 AA.
AC Q958U8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Cytochrome b.
GN CYTB.
OS Echinococcus granulosus.
OG Mitochondrion.
OC Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;
OC Cyclophyllidae; Taeniidae; Echinococcus.
OX NCBI_TaxID=6210;
RN [1]
RP SEQUENCE FROM N.A.
RA Le T.H., Blair D., Pearson M., McManus D.P.;
RT "Echinococcus mitochondrial genome."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
DR EMBL; AF297617; AAK51678.1; -.
DR InterPro; IPR001751; CaBP.s100.
DR Pfam; PF00032; cytochrome_b_c_1.
DR Pfam; PF00033; cytochrome_b_n; 1.
DR PROSITE; PS00193; CYTOCHROME_B_QQ; UNKNOWN_1.
DR PROSITE; PS00303; S100_CaBP; UNKNOWN_1.
DR Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;
KW Electron transport; Heme; Mitochondrion; Respiratory chain;
KW Transmembrane.
SQ SEQUENCE 355 AA; 41252 MW; D44ACFD58CAE84B7 CRC64;

Query Match 28.2%; Score 51; DB 8; Length 355;
Best Local Similarity 52.9%; Pred. No. 17;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 7 SKFIGITELGGCHGSEP 23
DB 317 SLFLSLTYLGGCHPEYP 333
| | : | | | | | |
| | : | | | | | |

RESULT 14
OY53N0 PRELIMINARY; PRT; 355 AA.
AC Q953N0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Cytochrome b.
GN CYTB.
OS Echinococcus granulosus.
OG Mitochondrion.
OC Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;
OC Cyclophyllidae; Taeniidae; Echinococcus.
OX NCBI_TaxID=6210;
RN [1]
RP SEQUENCE FROM N.A.
RA Le T.H., Dai T.H.N., Blair D., McManus D.P.;
RT "Complete mitochondrial genome of Echinococcus granulosus (G4 genotype)."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
DR EMBL; AF346403; AAK82343.1; -.
DR InterPro; IPR000179; Cyt_b_b6.
DR Pfam; PF00032; cytochrome_b_c_1.
DR Pfam; PF00033; cytochrome_b_n; 1.
DR PROSITE; PS00193; CYTOCHROME_B_QQ; UNKNOWN_1.
DR Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;
KW Electron transport; Heme; Mitochondrion; Respiratory chain;
KW Transmembrane.
SQ SEQUENCE 355 AA; 41220 MW; 61D849BF10DAED00 CRC64;

Query Match 28.2%; Score 51; DB 8; Length 355;
Best Local Similarity 52.9%; Pred. No. 17;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 7 SKFIGITELGGCHGSEP 23
DB 317 SLFLSLTYLGGCHPEYP 333
| | : | | | | | |
| | : | | | | | |

RESULT 15
OY1585 PRELIMINARY; PRT; 60 AA.
AC Q31585;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE (DB02) MHC class II beta 1 (Fragment).
OS Salmo salar (Atlantic salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX NCBI_TaxID=8030;
RN [1]
RP SEQUENCE FROM N.A.
RA Grimholt U., Olsaker I., de Vries Lindstrom C., Lie O.;
RT "A study of polymorphism in the MHC class II beta 1 and MHC class I
alpha 2 domain exons of Atlantic salmon (Salmo salar).";
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 12:49:05 : Search time 15.8667 Seconds  
(without alignments)  
268.741 Million cell updates/sec

Title: US-09-362-731A-5

Perfect score: 181

Sequence: 1 QYIKANSKFIGITELGGCHSEPCNIHRGKPF 32

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

1:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SID22/qcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SID22/qcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SID22/qcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23:	/SID22/qcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	181	100.0	32	21 AAY82636	Tetanus toxoid T C
2	160	88.4	137	21 AAY82634	Tetanus toxoid T C
3	98	54.1	31	21 AAY82632	Tetanus toxoid T C
4	92.5	51.1	92	19 AAW72338	Dermatophagoides D
5	92.5	51.1	92	20 AAY50511	Dermatophagoides S
6	92.5	51.1	92	22 AAU19114	House dust mite al
7	92	50.8	24	15 AAR51764	Der p II derived p
8	92	50.8	25	14 AAR36416	DPII-2(11-35) a De
9	92	50.8	25	16 AAR2550	Dermatophagoides p
10	92	50.8	25	19 AAW71939	Dermatophagoides D

11	92	50.8	25	20 AAY50393	Dermatophagoides s
12	92	50.8	25	22 AAU18996	T-cell epitope con
13	92	50.8	47	15 AAR62723	LHRH-containing im
14	92	50.8	61	19 AAW72340	Dermatophagoides D
15	92	50.8	61	20 AAY50513	Dermatophagoides s
16	92	50.8	61	22 AAU19116	House dust mite al
17	92	50.8	84	19 AAW72339	Dermatophagoides D
18	92	50.8	84	20 AAY50512	Dermatophagoides s
19	92	50.8	84	22 AAU19115	House dust mite al
20	92	50.8	129	15 AAR49921	Protein allergen o
21	92	50.8	129	15 AAR49922	Protein allergen o
22	92	50.8	129	19 AAW72335	Dermatophagoides D
23	92	50.8	129	19 AAW72337	Dermatophagoides D
24	92	50.8	129	20 AAY50508	Dermatophagoides s
25	92	50.8	129	20 AAY50510	Dermatophagoides s
26	92	50.8	129	22 AAU80061	Modified tick alle
27	92	50.8	129	22 AAG78996	House dust mite al
28	92	50.8	129	22 AAU07751	House dust mite al
29	92	50.8	129	22 AAU19111	House dust mite al
30	92	50.8	129	22 AAU19113	House dust mite al
31	92	50.8	129	23 ABG66971	House dust mite al
32	92	50.8	129	23 ABG66972	House dust mite al
33	92	50.8	129	23 ABG66974	House dust mite al
34	92	50.8	129	23 ABG66975	House dust mite al
35	92	50.8	129	23 ABG66976	House dust mite al
36	92	50.8	129	23 ABG66991	House dust mite al
37	92	50.8	129	23 ABG66992	House dust mite al
38	92	50.8	129	23 ABG66993	House dust mite al
39	92	50.8	129	23 ABG66994	House dust mite al
40	92	50.8	129	23 ABG66995	House dust mite al
41	92	50.8	129	23 ABG66996	House dust mite al
42	92	50.8	129	23 ABG66997	House dust mite al
43	92	50.8	129	23 ABG66998	House dust mite al
44	92	50.8	129	23 ABG66999	House dust mite al
45	92	50.8	129	23 ABG67000	House dust mite al

#### ALIGNMENTS

##### RESULT 1

AA82636  
ID AAY82636 standard; peptide; 32 AA.

XX AAY82636;

AC AAY82636;

DT 07-AUG-2000 (first entry)

XX Tetanus toxoid T cell epitope and Der pII B cell epitope peptide.

DE T cell epitope; B cell epitope; allergen; antigenic;  
KW antiallergic; antiasthmatic; antiinflammatory; dermatological;  
KW immunosuppressive; vaccine; rhinitis; sinusitis; bronchial asthma;  
KW atopic dermatitis; acute urticaria; chronic urticaria;  
KW gastro-intestinal syndrome; food allergen; oro-pharyngeal syndrome;  
KW anaphylactic reaction; drug hypersensitivity; allergic reaction.

XX Dermatophagoides pteronyssinus.

OS Clostridium tetani.

OS Synthetic.

XX WO200006694-A2.

XX 10-FEB-2000.

XX 20-JUL-1999; 99WO-BE00092.

XX 30-JUL-1998; 98EP-0870167.

XX (UNIO ) UCB SA.

XX Saint-Remy J, Jacquemin M;

XX

DR WPI; 2000-422470/36.  
 XX New compound for prevention and treatment of allergies comprises at  
 PT least one allergen antigenic determinant recognized by a B cell and at  
 PT least one antigenic determinant which does not trigger T cell  
 PT activation -  
 XX  
 PS Claim 8; Page 35; 50pp; English.  
 XX  
 CC The present invention describes a compound (I) for the prevention and/or  
 CC treatment of allergy. The compound comprises at least one allergen  
 CC antigenic determinant (i) recognised by a B cell or an antibody secreted  
 CC by a B cell of a non-atopic individual and at least one antigenic  
 CC determinant (ii) different from the allergen that triggers T cell  
 CC activation. (I) has anti-allergic, antiasthmatic, antiinflammatory,  
 CC dermatological and immunosuppressive activities, and can be used in a  
 CC vaccine. (I) may be used in a pharmaceutical or cosmetic medicament to  
 CC treat and/or prevent allergies or a disease of allergic origin,  
 CC especially hypersensitivities. These include rhinitis, sinusitis,  
 CC bronchial asthma, atopic dermatitis, some forms of acute and chronic  
 CC urticaria, gastro-intestinal syndromes associated with the ingestion of  
 CC food allergens, oro-pharyngeal syndrome, anaphylactic reactions  
 CC associated with drug hypersensitivities and/or a mixture of these. The  
 CC use of (I) in the treatment of allergic conditions avoids the need for  
 CC drug treatment, which often causes undesirable side-effects. Also, prior  
 CC art drug therapies alleviate symptoms, but do not influence their  
 CC causes, however (I) actually combats the cause of an allergic reaction.  
 CC The present sequence represents a specifically claimed compound peptide  
 CC sequence from the present invention.  
 XX  
 SQ Sequence 32 AA;  
 Query Match 100.0%; Score 181; DB 21; Length 32;  
 Best Local Similarity 100.0%; Pred. No. 6e-19;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 QYIKANSFIGITELGGCHGSEPCNIHRGKPF 32  
 |||||  
 Db 1 QYIKANSFIGITELGGCHGSEPCNIHRGKPF 32  
 |||||  
 RESULT 2  
 AAY82634  
 ID AAY82634 standard; peptide; 137 AA.  
 XX  
 AC AAY82634;  
 XX  
 DT 07-AUG-2000 (first entry)  
 XX  
 DE Tetanus toxoid T cell epitopes and Der pII B cell epitopes peptide.  
 XX  
 KW T cell epitope; B cell epitope; allergy; allergen; antigenic;  
 KW anti-allergic; antiasthmatic; antiinflammatory; dermatological;  
 KW immunosuppressive; vaccine; rhinitis; sinusitis; bronchial asthma;  
 KW atopic dermatitis; acute urticaria; chronic urticaria;  
 KW gastro-intestinal syndrome; food allergen; oro-pharyngeal syndrome;  
 KW anaphylactic reaction; drug hypersensitivity; allergic reaction.  
 XX  
 OS Dermatophagoides pteronyssinus.  
 OS Clostridium tetani.  
 OS Synthetic.  
 XX WO200006694-A2.  
 XX 10-FEB-2000.  
 XX  
 PD 10-FEB-2000.  
 XX  
 PF 20-JUL-1999; 99WO-BE00092.  
 XX  
 PR 30-JUL-1998; 98EP-0870167.  
 XX  
 PA (UNTO ) UCB SA.  
 XX  
 PI Saint-Remy J, Jacquemin M;

XX WPI; 2000-422470/36.  
 DR New compound for prevention and treatment of allergies comprises at  
 XX least one allergen antigenic determinant recognized by a B cell and at  
 PT least one antigenic determinant which does not trigger T cell  
 PT activation -  
 XX  
 PS Claim 8; Page 35; 50pp; English.  
 XX  
 CC The present invention describes a compound (I) for the prevention and/or  
 CC treatment of allergy. The compound comprises at least one allergen  
 CC antigenic determinant (i) recognised by a B cell or an antibody secreted  
 CC by a B cell of a non-atopic individual and at least one antigenic  
 CC determinant (ii) different from the allergen that triggers T cell  
 CC activation. (I) has anti-allergic, antiasthmatic, antiinflammatory,  
 CC dermatological and immunosuppressive activities, and can be used in a  
 CC vaccine. (I) may be used in a pharmaceutical or cosmetic medicament to  
 CC treat and/or prevent allergies or a disease of allergic origin,  
 CC especially hypersensitivities. These include rhinitis, sinusitis,  
 CC bronchial asthma, atopic dermatitis, some forms of acute and chronic  
 CC urticaria, gastro-intestinal syndromes associated with the ingestion of  
 CC food allergens, oro-pharyngeal syndrome, anaphylactic reactions  
 CC associated with drug hypersensitivities and/or a mixture of these. The  
 CC use of (I) in the treatment of allergic conditions avoids the need for  
 CC drug treatment, which often causes undesirable side-effects. Also, prior  
 CC art drug therapies alleviate symptoms, but do not influence their  
 CC causes, however (I) actually combats the cause of an allergic reaction.  
 CC The present sequence represents a specifically claimed compound peptide  
 CC sequence from the present invention.  
 XX  
 SQ Sequence 137 AA;  
 Query Match 88.4%; Score 160; DB 21; Length 137;  
 Best Local Similarity 90.6%; Pred. No. 2.9e-15;  
 Matches 29; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 Qy 1 QYIKANSFIGITELGGCHGSEPCNIHRGKPF 32  
 |||||  
 Db 19 QYIKANSFIGITELSSCHGSEPCIIHRGKPF 50  
 |||||  
 RESULT 3  
 AAY82632  
 ID AAY82632 standard; peptide; 31 AA.  
 XX  
 AC AAY82632;  
 XX  
 DT 07-AUG-2000 (first entry)  
 XX  
 DE Tetanus toxoid T cell epitope and Der pII B cell epitope peptide.  
 XX  
 KW T cell epitope; B cell epitope; allergy; allergen; antigenic;  
 KW anti-allergic; antiasthmatic; antiinflammatory; dermatological;  
 KW immunosuppressive; vaccine; rhinitis; sinusitis; bronchial asthma;  
 KW atopic dermatitis; acute urticaria; chronic urticaria;  
 KW gastro-intestinal syndrome; food allergen; oro-pharyngeal syndrome;  
 KW anaphylactic reaction; drug hypersensitivity; allergic reaction.  
 XX  
 OS Dermatophagoides pteronyssinus.  
 OS Clostridium tetani.  
 OS Synthetic.  
 XX WO200006694-A2.  
 XX 10-FEB-2000.  
 XX  
 PD 10-FEB-2000.  
 XX  
 PF 20-JUL-1999; 99WO-BE00092.  
 XX  
 PR 30-JUL-1998; 98EP-0870167.  
 XX  
 PA (UNTO ) UCB SA.  
 XX



CC may be used in (I). The sample and allergens are combined under  
CC conditions appropriate for the binding of blood components with the  
CC polypeptides. The extent of binding is then indicative of the  
CC sensitivity of the patient to house dust mites. (I) may be used to screen  
CC individuals for sensitivity to Dermatophagoides (house dust mites). The  
CC house dust mite is a major cause of a variety of allergic disorders such  
CC as asthma, rhinitis and eczematoid dermatitis. AAY50360-Y50342 and  
CC AAY50546-Y50555 represent house dust mite allergen peptide fragments  
CC derived from Der p I, Der f II, Der f I and Der f II.  
XX  
SQ Sequence 92 AA;

Query Match 51.1%; Score 92.5; DB 20; Length 92;  
Best Local Similarity 67.9%; Pred. No. 1e-05;  
Matches 19; Conservative 2; Mismatches 6; Indels 1; Gaps 1;  
QY 5 ANSKFIGITELGGCHGSEPCNIHRGKPF 32  
Db 9 ANSE-IKKVMVPGCHGSEPCIIHRGKPF 35  
III: I : ||||| |||||

RESULT 6  
AAU19114  
ID AAU19114 standard; Protein; 92 AA.  
XX  
AC AAU19114;  
XX  
DT 04-DEC-2001 (first entry)  
XX  
DE House dust mite allergen Der f II variant MT5.  
XX  
KW House dust mite; allergenic protein; Der p I; Der p II; Der f I;  
KW Der f II; antiallergenic; immunostimulant; house dust mite allergy;  
KW T-cell epitope.  
XX

OS Dermatophagoides farinae.  
XX  
XX US268491-B1.  
XX  
XX 31-JUL-2001.  
XX  
XX 07-JUN-1995; 95US-0484296.  
XX  
XX 19-MAY-1995; 95US-0445307.  
XX 16-OCT-1991; 91US-0777859.  
XX 08-MAY-1992; 92US-0881396.  
XX 14-APR-1993; 93WO-US03471.  
XX 14-APR-1994; 94US-0227772.  
XX  
XX (IMMU-) IMMULOGIC PHARM CORP.  
XX  
XX Garman RD, Greenstein JL, Kuo M, Rogers BL, Franzen HM, Chen X;  
XX Evans S, Shaked Z;  
XX  
XX WPI; 2001-549074/61.  
XX

PT Peptides comprising T cell groups of the major allergens from  
PT Dermatophagoides (house dust mites), useful for treating house dust  
PT mite allergy in humans, and for diagnosing sensitivity to house dust  
PT mite protein allergens -  
XX  
PS Disclosure; Figure 24; 158pp; English.

XX The invention relates to an isolated peptide of the major protein  
CC allergens of the genus Dermatophagoides, which comprises at least one T  
CC cell group of a protein allergen from Der p (DP) I, DP II, Der f (DF) I  
CC or DF II. The isolated peptide comprises at least two regions, an  
CC each region comprising at least one T cell group of a protein allergen  
CC of the genus Dermatophagoides. The regions are derived from the same or  
CC different protein allergens of the genus Dermatophagoides. The peptides  
CC are useful for treating house dust mite allergy in humans. The peptides  
CC are also useful for detecting or diagnosing sensitivity to house dust  
CC mite protein allergens. The present peptides have similar or enhanced

CC therapeutic properties as the naturally-occurring allergen, but have  
CC reduced side effects, and increased solubility and stability. The  
CC present sequence represents an allergenic protein from Dermatophagoides  
CC from which the T-cell epitope containing peptides are derived.  
XX

SQ Sequence 92 AA;  
Query Match 51.1%; Score 92.5; DB 22; Length 92;  
Best Local Similarity 67.9%; Pred. No. 1e-05;  
Matches 19; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 5 ANSKFIGITELGGCHGSEPCNIHRGKPF 32  
Db 9 ANSE-IKKVMVPGCHGSEPCIIHRGKPF 35  
III: I : ||||| |||||  
RESULT 7  
AAR51764  
ID AAR51764 standard; Protein; 24 AA.  
XX  
AC AAR51764;  
XX  
DT 01-FEB-1995 (first entry)  
XX  
DE Der p II derived peptide, DP II-2(11-35).  
XX

KW Group I; protein allergen; house dust mite; D. pteronyssinus; Der p I;  
KW homology; D. farinae; Der f I; group II; Der p II; Der f II; T-cell;  
KW epitopes; fusion peptides; antigenic fragments; substitution; deletion;  
KW addition; chemical synthesis; chemical cleavage; recombinant techniques;  
KW allergic response; immunoglobulin E; IgE; immunotherapy; anaphylaxis;  
KW IgE-mediated responses; anergise; lymphokine secretion profile; modify;  
KW T cell subpopulations; unresponsive; Immune response; tolerise.  
XX

OS Dermatophagoides pteronyssinus.  
XX  
XX ZA9302677-A.  
XX  
XX 26-JAN-1994.  
XX  
XX 16-APR-1993; 93ZA-0002677.  
XX  
XX 16-APR-1993; 93ZA-0002677.  
XX  
XX (IMMU-) IMMULOGIC PHARM CORP.  
XX  
XX Garman RD, Greenstein JL, Kuo M, Rogers BL;  
XX WPI; 1994-126807/15.  
XX

PT Isolated and/or modified peptides comprising T-cell epitopes - of  
PT major protein allergens of genus Dermatophagoides, used to treat  
PT or diagnose sensitivity to house dust mites  
XX  
PS Disclosure; Fig 3; 154pp; English.

XX The sequences given in AAR51731-841 represent T-cell epitopes derived  
CC from the group I and II protein allergens from the house dust mite D.  
CC farinae and D. pteronyssinus, Der f I, Der f II, Der p I and Der p II  
CC respectively. The Der f II proteinsCC shows high homology having an  
CC identity of 88%, with an identity of 81% between the two group I  
CC proteins (see also AAR51727-30). Fusion peptides may be produced which  
CC comprise at least two or these antigenic fragments. Each region of  
CC these fusion peptides may be derived from the same, or different, mite  
CC allergens. The antigenic fragments may be altered by substitution,  
CC deletion or addition to enhance their antigenicity. These peptides may  
CC be produced by chemical synthesis, chemical cleavage of the protein  
CC allergen or by recombinant techniques. These peptides, or the fusion  
CC peptides, when administered to a house dust mite sensitive individual,  
CC are capable of modifying the allergic response of the individual to the  
CC allergen. The peptides do not bind to immunoglobulin E (IgE), or bind  
CC IgE to a lesser extent than the full length protein allergen. This  
CC reduces the major complications of standard immunotherapy, which are



CC IgE-mediated responses such as anaphylaxis. Exposure of mite allergic  
 CC patients to these peptides may tolerate or anergise appropriate T cell  
 CC subpopulations such that they become unresponsive to mite allergens and  
 CC do not participate in mounting an immune response upon exposure.  
 CC Administration of the peptides may also modify the lymphokine secretion  
 CC profile as compared with exposure to the naturally occurring mite protein  
 CC allergen.

XX Sequence 24 AA;

Query Match 50.8%; Score 92; DB 15; Length 24;  
 Best Local Similarity 93.8%; Pred. No. 3e-06;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 17 GCHGSEPCNIHRGKPF 32

DB 9 GCHGSEPCNIHRGKPF 24

RESULT 8

ID AAR36416 standard; peptide; 25 AA.

XX AAR36416;

DT 12-AUG-1993 (first entry)

XX DPII-2(11-35) a Dermatophagoides protein allergen.

DE T cell epitope; house dust mite; allergy; soluble; Der pII.

KW Synthetic.

OS WO9308279-A.

PN 29-APR-1993.

PD 15-OCT-1992; 92WO-US08637.

PF 16-OCT-1991; 91US-0777859.

PR 08-MAY-1992; 92US-0881396.

XX (IMMU-) IMMULOGIC PHARM CORP.

PA Garman RD, Greenstein JL, Kuo MC, Rogers BL;

PI WPI; 1993-152472/18.

XX Isolated peptide(s) of Dermatophagoides protein allergens - for  
 PT diagnosis and treatment of sensitivity to house dust mite

XX Claim 44; Fig 3; 176pp; English.

XX The peptide is one of a series of overlapping peptides synthesised by  
 CC standard techniques to cover the whole Dermatophagoides  
 CC pteronyssinus Der pII sequence. The T cell epitopes of the protein  
 CC were mapped by detection of the peptide's ability to stimulate T  
 CC cell activity. The peptides may be used for diagnosis and treatment  
 CC of sensitivity to house dust mite allergens. When administered to  
 CC house dust mite sensitive individuals, the peptides are capable of  
 CC modifying the allergic response to the allergens. The peptides may  
 CC be modified for e.g. increasing solubility, enhancing therapeutic or  
 CC preventive efficacy or stability.  
 CC See also AAR34686-700 and AAR36398-490.

XX Sequence 25 AA;

Query Match 50.8%; Score 92; DB 14; Length 25;  
 Best Local Similarity 93.8%; Pred. No. 3.1e-06;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 17 GCHGSEPCNIHRGKPF 32

XXXXXXXXXXXXXXXXXXXX

DB 10 GCHGSEPCNIHRGKPF 25

RESULT 9

AAR82550

ID AAR82550 standard; peptide; 25 AA.

XX AAR82550;

AC 30-MAY-1996 (first entry)

DT Dermatophagoides pteronyssinus group II peptide DPII-1 (11-35).

DE House dust mite; DerpI; DerfII; DerfII; allergen; allergy;  
 KW overlapping peptide; screening.

XX Dermatophagoides pteronyssinus.

XX WO9528424-A1.

XX 26-OCT-1995.

XX 12-APR-1995; 95WO-US04481.

XX 14-APR-1994; 94US-0227722.

XX (IMMU-) IMMULOGIC PHARM CORP.

XX Chen X, Evans S, Franzen HM, Kuo M, Shaked Z;

XX WPI; 1995-373765/48.

XX Compens. contg. house mite allergen-derived peptide(s), some of  
 PT which are new - are used to treat allergy, and are stable, soluble  
 PT and able to induce T cell non-responsiveness

XX Disclosure; Fig 2; 61pp; English.

XX Claimed therapeutic compositions contain at least one of the peptides  
 CC DPI-21.2 and DPI-22.2 and also at least one of the new peptides  
 CC DPI-23.31, DPI-26.6, DPII-20.9, DPII-22.14 and DPII-25.15. The  
 CC compositions are useful for treating sensitivity to house dust mite  
 CC allergens. The peptides were identified by screening overlapping  
 CC peptides derived from D.pteronyssinus and D.farinae group I and II  
 CC allergens for T-cell reactivity in sensitised individuals.  
 CC The present sequence is that of overlapping peptide DPII-2 (11-35).

XX Sequence 25 AA;

Query Match 50.8%; Score 92; DB 16; Length 25;

Best Local Similarity 93.8%; Pred. No. 3.1e-06;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 17 GCHGSEPCNIHRGKPF 32

XXXXXXXXXXXXXXXXXXXX

RESULT 10

AAW71939

ID AAW71939 standard; peptide; 25 AA.

XX AAW71939;

DT 16-DEC-1998 (first entry)

DE Dermatophagoides Der p II protein peptide DPII-2.

KW genus Dermatophagoides; major protein allergen; T cell epitope;  
 KW Der p I; Der p II; Der f I; Der f II; house dust mite allergy.

XX Dermatophagoides sp.

```

PN US5820862-A.
XX
PD 13-OCT-1998.
XX
PF 07-JUN-1995; 95US-0482142.
XX
PR 19-MAY-1995; 95US-0445307.
PR 14-APR-1994; 94US-0227772.
PR 07-JUN-1995; 95US-0482142.
XX
PA (IMMU-) IMMULOGIC PHARM CORP.
XX
PI Chen X, Evans S, Franzen HM, Garman RD, Greenstein JL;
PI Kuo M, Rogers BL, Shaked Z;
XX
DR WPI; 1998-567590/48.
XX
PT Dermatophagoides allergen peptides - useful for treating house dust
PT mite allergy
XX
PS Disclosure; Column 99-100; 155pp; English.
XX
CC The present invention describes peptides for treating sensitivity to
CC house dust mite allergens from the genus Dermatophagoides. Peptides
CC within the scope of the invention comprise at least one T cell epitope,
CC or preferably at least two T cell epitopes of a protein allergen
CC selected from the allergens Der p I, Der p II, Der f I, or Der f II.
CC The invention also describes modified peptides having similar or
CC enhanced therapeutic properties as the corresponding, naturally
CC occurring allergen, but having reduced side effects. AAW71912 to
CC AAW72000, and AAW72257 to AAW72330 represent peptides from the present
CC invention.
XX
SQ Sequence 25 AA;

Query Match 50.8%; Score 92; DB 19; Length 25;
Best Local Similarity 93.8%; Pred. No. 3.1e-06;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 17 GCHGSEPCNIHRGKPF 32
Db 10 GCHGSEPCNIHRGKPF 25

RESULT 11
AAV50393
ID AAV50393 standard; Peptide; 25 AA.
XX
AC AAV50393;
XX
DT 25-JAN-2000 (first entry)
XX
DE Dermatophagoides sp major protein allergen DP II-2.
XX
KW Allergen; house dust mite; detection; sensitivity; T cell epitope;
KW screening; allergic disorder; asthma; rhinitis; ectopic dermatitis;
KW Der f I; Der p I; Der p II; Der f II.
XX
OS Dermatophagoides sp.
XX
PN US5968526-A.
XX
PD 19-OCT-1999.
XX
PF 07-JUN-1995; 95US-0478572.
XX
PR 19-MAY-1995; 95US-0445307.
PR 14-APR-1994; 94US-0227772.
PR 12-APR-1995; 95WO-US04481.
XX
PA (IMMU-) IMMULOGIC PHARM CORP.
XX
PI Garman RD, Greenstein JL, Rogers BL, Franzen HM, Shaked Z, Chen X;

```

```

PI Evans S, Kuo M;
XX
DR WPI; 1999-590385/50.
XX
PT Screening individuals for allergic reactions to T cell epitopes of
PT major allergens from house dust mites -
XX
PS Disclosure; Column 101-102; 158pp; English.
XX
CC This invention describes a novel method (I) for detecting whether an
CC individual is sensitive to Dermatophagoides (house dust mites). The
CC method involves detecting sensitivity to house dust mites in patients,
CC comprising combining a blood sample from the individual with 1 or more
CC isolated T cell epitopes of the protein allergens I and II (DP I) and
CC (DP II) from Dermatophagoides (house dust mites). 32 T cell epitopes
CC with varying, defined amino acids sequences (given in the specification)
CC may be used in (I). The sample and allergens are combined under the
CC conditions appropriate for the binding of blood components with the
CC polypeptides. The extent of binding is then indicative of the
CC sensitivity of the patient to house dust mites. (I) may be used to screen
CC individuals for sensitivity to Dermatophagoides (house dust mites). The
CC house dust mite is a major cause of a variety of allergic disorders such
CC as asthma, rhinitis and ectopic dermatitis. AAY50360-Y50542 and
CC AAY50546-Y50555 represent house dust mite allergen peptide fragments
CC derived from Der p I, Der f II, Der f I and Der f II.
XX
SQ Sequence 25 AA;

Query Match 50.8%; Score 92; DB 20; Length 25;
Best Local Similarity 93.8%; Pred. No. 3.1e-06;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 17 GCHGSEPCNIHRGKPF 32
Db 10 GCHGSEPCNIHRGKPF 25

RESULT 12
AAU18996
ID AAU18996 standard; Peptide; 25 AA.
XX
AC AAU18996;
XX
DT 04-DEC-2001 (first entry)
XX
DE T-cell epitope containing peptide DP11-2.
XX
KW House dust mite; allergenic peptide; Der p I; Der p II; Der f I;
KW Der f II; antiallergenic; immunostimulant; house dust mite allergy;
KW T-cell epitope.
XX
OS Dermatophagoides pteronyssinus.
XX
PN US6268491-B1.
XX
PD 31-JUL-2001.
XX
PF 07-JUN-1995; 95US-0484296.
XX
PR 19-MAY-1995; 95US-0445307.
PR 16-OCT-1991; 91US-0777859.
PR 08-MAY-1992; 92US-0881396.
PR 14-APR-1993; 93WO-US03471.
PR 14-APR-1994; 94US-0227772.
XX
PA (IMMU-) IMMULOGIC PHARM CORP.
XX
PI Garman RD, Greenstein JL, Kuo M, Rogers BL, Franzen HM, Chen X;
PI Evans S, Shaked Z;
XX
DR WPI; 2001-549074/61.
XX
PT Peptides comprising T cell groups of the major allergens from

```

PT Dermatophagoides (house dust mites), useful for treating house dust  
PT mite allergy in humans, and for diagnosing sensitivity to house dust  
PT mite protein allergens -  
XX  
XX  
PS Claim 5; Figure 3; 158pp; English.  
XX  
XX The invention relates to an isolated peptide of the major protein  
CC allergens of the genus Dermatophagoides, which comprises at least one T  
CC cell group of a protein allergen from Der p (DP) I, DP II, Der f (DF) I  
CC or DF II. The isolated peptide comprises at least two regions,  
CC each region comprising at least one T cell group of a protein allergen  
CC of the genus Dermatophagoides. The regions are derived from the same or  
CC different protein allergens of the genus Dermatophagoides. The peptides  
CC are useful for treating house dust mite allergy in humans. The peptides  
CC are also useful for detecting or diagnosing sensitivity to house dust  
CC mite protein allergens. The present peptides have similar or enhanced  
CC therapeutic properties as the naturally-occurring allergen, but have  
CC reduced side effects, and increased solubility and stability. The  
CC present sequence represents an allergenic T-cell epitope containing  
CC peptide derived from the Dermatophagoides allergenic proteins.  
XX  
SQ Sequence 25 AA;  
Query Match 50.8%; Score 92; DB 22; Length 25;  
Best Local Similarity 93.8%; Pred. No. 3.1e-06;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 17 GCHGSEPCNIHRGKPF 32  
||||||| |||||  
DB 10 GCHGSEPCIIHRGKPF 25  
RESULT 13  
AAR62723  
ID AAR62723 standard; peptide; 47 AA.  
XX  
AC AAR62723;  
XX  
DT 17-SEP-1995 (first entry)  
XX  
DE LHRH-containing immunogenic peptide.  
XX  
KW Helper T cell epitope; universal immune stimulator; invasive; haptens;  
KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;  
KW androgen-dependent carcinoma; antitumour; infertility;  
KW tetanus toxin.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Domain 1..16  
FT /note= "invasin domain"  
FT Domain 19..35  
FT /note= "tetanus toxin helper T cell epitope"  
FT Domain 38..47  
FT /note= "LHRH haptens"  
XX  
XX WO9425060-A.  
XX  
XX 10-NOV-1994.  
XX  
XX 28-APR-1994; 94WO-US04832.  
XX  
XX 27-APR-1993; 93US-0057166.  
XX 14-APR-1994; 94US-0229275.  
XX  
XX (LADD/) LADD A E.  
XX (WANG/) WANG C Y.  
XX (ZAMB/) ZAMB T.  
XX  
XX Ladd AE, Wang CY, Zamb T;  
XX  
XX WPI; 1994-357910/44.  
XX

XX Immunogenic luteinising hormone releasing hormone peptide(s) -  
PT that suppress LHRH activity in males and females  
XX  
XX Claim 8; Page 88; 213pp; English.  
XX  
XX Synthetic immunogenic peptides are provided in which a universal immune  
CC stimulator is linked to a peptide or protein haptens containing B cell  
CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes  
CC potent immune responses to the coupled peptide or protein. The  
CC stimulator consists of (A) a promiscuous helper T cell epitope (Th)  
CC which elicits an immune response to the coupled peptide in members of  
CC a heterogeneous population expressing diverse HLA phenotypes, and (B)  
CC an adjuvant peptide sequence from the invasive protein of Yersinia.  
CC Spacer amino acid sequences (e.g. Gly-Gly) can be provided between the  
CC invasive and Th domains and between the immune stimulator and haptens  
CC components. When the haptens is LHRH, then optionally the invasive domain  
CC can be omitted from the immune stimulator component.  
CC The present sequence represents an LHRH-containing immunogenic peptide  
CC as above which can be used as a potent vaccine for treating e.g.  
CC prostatic hyperplasia, androgen-dependent carcinoma, prostatic  
CC carcinoma, testicular carcinoma, endometriosis, benign uterine tumours,  
CC recurrent functional ovarian cysts, (severe) premenstrual syndrome or  
CC oestrogen-dependent breast cancer, or for induction of infertility.  
XX  
SQ Sequence 47 AA;  
Query Match 50.8%; Score 92; DB 15; Length 47;  
Best Local Similarity 90.5%; Pred. No. 6e-06;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 QYIKANSFIGITELGGCHGS 21  
||||||| ||||| ||  
DB 21 QYIKANSFIGITELGGHWS 41  
RESULT 14  
AAW72340  
ID AAW72340 standard; peptide; 61 AA.  
XX  
AC AAW72340;  
XX  
DT 16-DEC-1998 (first entry)  
XX  
DE Dermatophagoides Der f II clone MTL6.  
XX  
KW genus Dermatophagoides; major protein allergen; T cell epitope;  
KW Der p I; Der p II; Der f I; Der f II; house dust mite allergy.  
XX  
OS Dermatophagoides sp.  
XX  
XX US5820862-A.  
XX  
XX 13-OCT-1998.  
XX  
XX 07-JUN-1995; 95US-0482142.  
XX  
XX 19-MAY-1995; 95US-0445307.  
XX 14-APR-1994; 94US-0227772.  
XX 07-JUN-1995; 95US-0482142.  
XX  
XX (IMMU-) IMMULOGIC PHARM CORP.  
XX  
XX Chen X, Evans S, Franzen HM, Garman RD, Greenstein JL;  
XX Kuo M, Rogers BL, Shaked Z;  
XX  
XX WPI; 1998-567590/48.  
XX  
XX Dermatophagoides allergen peptides - useful for treating house dust  
XX mite allergy  
XX  
XX Disclosure; Fig 24; 155pp; English.  
XX

CC The present invention describes peptides for treating sensitivity to  
 CC house dust mite allergens from the genus Dermatophagoides. Peptides  
 CC within the scope of the invention comprise at least one T cell epitope,  
 CC or preferably at least two T cell epitopes of a protein allergen  
 CC selected from the allergens Der p I, Der p II, Der f I, or Der f II.  
 CC The invention also describes modified peptides having similar or  
 CC enhanced therapeutic properties as the corresponding, naturally  
 CC occurring allergen, but having reduced side effects. The present  
 CC sequence represents a Der f II clone from the present invention.  
 XX  
 SQ Sequence 61 AA;  
 Query Match 50.8%; Score 92; DB 19; Length 61;  
 Best Local Similarity 93.8%; Pred. No. 7.9e-06;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 17 GCHGSEPCNIHRGKPF 32  
 ||||| |||||  
 Db 20 GCHGSEPCIIHRGKPF 35  
 ||||| |||||  
 RESULT 15  
 AAY50513  
 ID AAY50513 standard; Peptide; 61 AA.  
 XX  
 AC AAY50513;  
 XX  
 DT 25-JAN-2000 (first entry)  
 XX  
 DE Dermatophagoides sp major protein allergen fragment 13.  
 XX  
 KW Allergen; house dust mite; detection; sensitivity; T cell epitope;  
 KW screening; allergic disorder; asthma; rhinitis; ectopic dermatitis;  
 KW Der f I; Der p I; Der p II; Der f II.  
 XX  
 OS Dermatophagoides sp.  
 XX  
 PN US5968526-A.  
 XX  
 PD 19-OCT-1999.  
 XX  
 PF 07-JUN-1995; 95US-0478572.  
 XX  
 PR 19-MAY-1995; 95US-0445307.  
 PR 14-APR-1994; 94US-0227772.  
 PR 12-APR-1995; 95WO-US04481.  
 XX  
 PA (IMMU-) IMMULOGIC PHARM CORP.  
 XX  
 PI Garman RD, Greenstein JL, Rogers BL, Franzen HM, Shaked Z, Chen X;  
 PI Evans S, Kuo M;  
 XX  
 DR WPI; 1999-590385/50.  
 XX  
 PT Screening individuals for allergic reactions to T cell epitopes of  
 PT major allergens from house dust mites -  
 XX  
 PS Disclosure; Column 161-162; 158pp; English.  
 XX  
 CC This invention describes a novel method (I) for detecting whether an  
 CC individual is sensitive to Dermatophagoides (house dust mites). The  
 CC method involves detecting sensitivity to house dust mites in patients,  
 CC comprising combining a blood sample from the individual with 1 or more  
 CC isolated T cell epitopes of the protein allergens I and II (DP I) and  
 CC (DP II)) from Dermatophagoides (house dust mites). 32 T cell epitopes  
 CC with varying, defined amino acids sequences (given in the specification  
 CC may be used in (I). The sample and allergens are combined under  
 CC conditions appropriate for the binding of blood components with the  
 CC polypeptides. The extent of binding is then indicative of the  
 CC sensitivity of the patient to house dust mites. (I) may be used to screen  
 CC individuals for sensitivity to Dermatophagoides (house dust mites). The  
 CC house dust mite is a major cause of a variety of allergic disorders such  
 CC as asthma, rhinitis and ectopic dermatitis. AAY50360-Y50542 and

CC AAY50546-Y50555 represent house dust mite allergen peptide fragments  
 CC derived from Der p I, Der f II, Der f I and Der f II.  
 XX  
 SQ Sequence 61 AA;  
 Query Match 50.8%; Score 92; DB 20; Length 61;  
 Best Local Similarity 93.8%; Pred. No. 7.9e-06;  
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 17 GCHGSEPCNIHRGKPF 32  
 ||||| |||||  
 Db 20 GCHGSEPCIIHRGKPF 35  
 ||||| |||||  
 Search completed: December 4, 2002, 12:51:16  
 Job time : 16.1167 secs



```
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-09-949-889-4

Query Match      47.5%; Score 86; DB 10; Length 129;
Best Local Similarity 87.5%; Pred. No. 1e-05;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 17 GCGSEPCNIHRGKPF 32
      ||||| |||||
Db 20 GCGSEPCIIHRGAPF 35

RESULT 3
US-09-848-834A-11
; Sequence 11, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Apton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 830-844 of the
; OTHER INFORMATION: Tetanus toxoid precursor (Tentoxylisin) linked by a spacer to am
; OTHER INFORMATION: ino acid sequence 2-10 of the GnRH hormone
; NAME/KEY: MOD.RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Amidated-glutamine
; NAME/KEY: MOD.RES
; LOCATION: (28)..(28)
; OTHER INFORMATION: Amidated-glycine or glycineamide
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(15)
; OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus Toxoid Precursor
; OTHER INFORMATION: (Tentoxylisin)
; NAME/KEY: PEPTIDE
; LOCATION: (16)..(19)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (20)..(28)
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-11

Query Match      44.2%; Score 80; DB 10; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITELG 16
      ||||| ||||| |||||
Db 1 QYIKANSKFIGITELG 16

RESULT 4
US-09-943-548-2
; Sequence 2, Application US/09943548
; Patent No. US20020042364A1
; GENERAL INFORMATION:
; APPLICANT: Rittershaus, Charles W.
; TITLE OF INVENTION: MODULATION OF CHOLESTERYL ESTER TRANSFER PROTEIN (CETP) ACTIVITY
; FILE REFERENCE: TCS-411.1P US-1; TCS-411.1P US-2
; CURRENT APPLICATION NUMBER: US/09/943,548
```

```
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 08/432,483
; PRIOR FILING DATE: 1995-05-01
; PRIOR APPLICATION NUMBER: PCT/US96/06147
; PRIOR FILING DATE: 1996-05-01
; PRIOR APPLICATION NUMBER: 08/945,289
; PRIOR FILING DATE: 1997-10-17
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: vaccine peptide of the invention
US-09-943-548-2

Query Match      42.0%; Score 76; DB 10; Length 31;
Best Local Similarity 93.8%; Pred. No. 6.5e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITELG 16
      ||||| ||||| |||||
Db 2 QYIKANSKFIGITEFG 17

RESULT 5
US-09-862-849-2
; Sequence 2, Application US/09862849
; Patent No. US20020013274A1
; GENERAL INFORMATION:
; APPLICANT: Sudhir Paul
; APPLICANT: Larry J. Smith
; APPLICANT: Gennady Gololobov
; TITLE OF INVENTION: Methods for Identifying Inducers and Inhibitors of Proteolytic
; FILE REFERENCE: Antibodies, Compositions and Their Uses
; FILE REFERENCE: UNMC 63123 DIV
; CURRENT APPLICATION NUMBER: US/09/862,849
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: US 09/046,373
; PRIOR FILING DATE: 1998-03-23
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Clostridium tetani
US-09-862-849-2

Query Match      40.9%; Score 74; DB 10; Length 15;
Best Local Similarity 100.0%; Pred. No. 5.9e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15
      ||||| ||||| |||||
Db 1 QYIKANSKFIGITEL 15

RESULT 6
US-09-848-834A-2
; Sequence 2, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Apton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
```

; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Tetanus bacillus  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(16)  
; OTHER INFORMATION: Amino acid sequence 829-844 of the Tetanus  
; OTHER INFORMATION: Toxoid Precursor (Tentoxylisin)  
US-09-848-834A-2

Query Match 40.9%; Score 74; DB 10; Length 16;  
Best Local Similarity 100.0%; Pred. No. 6.3e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15  
| | | | | | | | | | | | | | | |  
Db 2 QYIKANSKFIGITEL 16

RESULT 7  
US-09-732-754-1  
; Sequence 1, Application US/09732754  
; Patent No. US20020031523A1  
; GENERAL INFORMATION:  
; APPLICANT: DRUILHE, PIERRE  
; APPLICANT: GRAS-MASSE, HELENE  
; APPLICANT: BENMORAMED, ELACHIR  
; TITLE OF INVENTION: SYSTEMIC IMMUNE RESPONSE INDUCED BY MUCOSAL ADMINISTRATION OF LIP  
; TITLE OF INVENTION: TAILED POLYPEPTIDES WITHOUT ADJUVANT  
; FILE REFERENCE: 200805US55  
; CURRENT APPLICATION NUMBER: US/09/732,754  
; CURRENT FILING DATE: 2000-12-11  
; PRIOR APPLICATION NUMBER: US 60/169,952  
; PRIOR FILING DATE: 1999-12-09  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 29  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Polypeptide  
; NAME/KEY: LIPID  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: N-epsilon Pam  
US-09-732-754-1

Query Match 40.9%; Score 74; DB 10; Length 29;  
Best Local Similarity 93.8%; Pred. No. 0.00012;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITELG 16  
| | | | | | | | | | | | | | | |  
Db 4 QYIKANSKFIGITERG 19

RESULT 8  
us-09-848-834A-15  
; Sequence 15, Application US/09848834A  
; Patent No. US20020076416A1  
; GENERAL INFORMATION:  
; APPLICANT: Aphton Corporation  
; TITLE OF INVENTION: Chimeric Peptide Immunogens  
; FILE REFERENCE: 1102865-0047  
; CURRENT APPLICATION NUMBER: US/09/848,834A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/202,328  
; PRIOR FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 15  
; LENGTH: 31  
; TYPE: PRT

; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the  
; OTHER INFORMATION: RH hormone linked by a spacer to amino sequence 830-844 of the  
; OTHER INFORMATION: tanus toxoid precursor (Tentoxylisin)  
; NAME/KEY: MOD\_RES  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(10)  
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone  
; NAME/KEY: PEPTIDE  
; LOCATION: (11)..(16)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (17)..(31)  
; OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus toxoid precursor  
; OTHER INFORMATION: (Tentoxylisin)  
US-09-848-834A-15

Query Match 40.9%; Score 74; DB 10; Length 31;  
Best Local Similarity 100.0%; Pred. No. 0.00013;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15  
| | | | | | | | | | | | | | | |  
Db 17 QYIKANSKFIGITEL 31

RESULT 9  
US-09-848-834A-19  
; Sequence 19, Application US/09848834A  
; Patent No. US20020076416A1  
; GENERAL INFORMATION:  
; APPLICANT: Aphton Corporation  
; TITLE OF INVENTION: Chimeric Peptide Immunogens  
; FILE REFERENCE: 1102865-0047  
; CURRENT APPLICATION NUMBER: US/09/848,834A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 60/202,328  
; PRIOR FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 19  
; LENGTH: 46  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of hum  
; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 830-844 of Teta  
; OTHER INFORMATION: toxoid precursor (Tentoxylisin) linked by a spacer to amino ac  
; OTHER INFORMATION: sequence 1-10 of GnRH  
; NAME/KEY: MOD\_RES  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline  
; NAME/KEY: MOD\_RES  
; LOCATION: (46)..(46)  
; OTHER INFORMATION: Amidated glycine or glycineamide  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(10)  
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone  
; NAME/KEY: PEPTIDE  
; LOCATION: (11)..(16)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (17)..(31)  
; OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus toxoid precursor.  
; OTHER INFORMATION: (Tentoxylisin)  
; NAME/KEY: PEPTIDE  
; LOCATION: (32)..(37)  
; OTHER INFORMATION: Spacer peptide  
; NAME/KEY: PEPTIDE  
; LOCATION: (38)..(46)

OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone  
US-09-848-834A-19

Query Match 40.9%; Score 74; DB 10; Length 46;  
Best Local Similarity 100.0%; Pred. No. 0.00019;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15  
|||||

Db 17 QYIKANSKFIGITEL 31

## RESULT 10

US-09-943-548-8  
; Sequence 8, Application US/09943548  
; Patent No. US20020042364A1  
; GENERAL INFORMATION:  
; APPLICANT: Rittershaus, Charles W.  
; APPLICANT: Thomas, Lawrence J.  
; TITLE OF INVENTION: MODULATION OF CHOLESTERYL ESTER TRANSFER PROTEIN (CETP) ACTIVITY  
; FILE REFERENCE: TCS-411.1P US-1; TCS-411.1P US-2  
; CURRENT APPLICATION NUMBER: US/09/943,548  
; CURRENT FILING DATE: 2001-08-30  
; PRIOR APPLICATION NUMBER: 08/432,483  
; PRIOR FILING DATE: 1995-05-01  
; PRIOR APPLICATION NUMBER: PCT/US96/06147  
; PRIOR FILING DATE: 1996-05-01  
; PRIOR APPLICATION NUMBER: 08/945,289  
; PRIOR FILING DATE: 1997-10-17  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8  
; LENGTH: 50  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: vaccine peptide of the invention  
US-09-943-548-8

Query Match 40.9%; Score 74; DB 10; Length 50;

Best Local Similarity 100.0%; Pred. No. 0.00021;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITEL 15  
|||||

Db 2 QYIKANSKFIGITEL 16

## RESULT 11

US-09-983-019-8  
; Sequence 8, Application US/09983019  
; Patent No. US20020146820A1  
; GENERAL INFORMATION:  
; APPLICANT: Diamond, Don J.  
; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS P  
; FILE REFERENCE: 1954-347US  
; CURRENT APPLICATION NUMBER: US/09/983,019  
; CURRENT FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: US 60/241,944  
; PRIOR FILING DATE: 2000-10-20  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 8  
; LENGTH: 29  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: {}...()  
; OTHER INFORMATION: Human cytomegalovirus vaccine peptide  
US-09-983-019-8

Query Match 38.7%; Score 70; DB 10; Length 29;

Best Local Similarity 100.0%; Pred. No. 0.00046;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITE 14  
|||||

Db 16 QYIKANSKFIGITE 29

## RESULT 12

US-09-983-019-9  
; Sequence 9, Application US/09983019  
; Patent No. US20020146820A1  
; GENERAL INFORMATION:  
; APPLICANT: Diamond, Don J.  
; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS P  
; FILE REFERENCE: 1954-347US  
; CURRENT APPLICATION NUMBER: US/09/983,019  
; CURRENT FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: US 60/241,944  
; PRIOR FILING DATE: 2000-10-20  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 9  
; LENGTH: 29  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: {}...()  
; OTHER INFORMATION: Human cytomegalovirus vaccine peptide  
US-09-983-019-9

Query Match 38.7%; Score 70; DB 10; Length 29;

Best Local Similarity 100.0%; Pred. No. 0.00046;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITE 14  
|||||

Db 1 QYIKANSKFIGITE 14

## RESULT 13

US-09-983-019-3  
; Sequence 3, Application US/09983019  
; Patent No. US20020146820A1  
; GENERAL INFORMATION:  
; APPLICANT: Diamond, Don J.  
; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS P  
; FILE REFERENCE: 1954-347US  
; CURRENT APPLICATION NUMBER: US/09/983,019  
; CURRENT FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: US 60/241,944  
; PRIOR FILING DATE: 2000-10-20  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 3  
; LENGTH: 31  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: {}...()  
; OTHER INFORMATION: Human cytomegalovirus vaccine peptide  
; NAME/KEY: LIPID  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: di-palmitic acid  
US-09-983-019-3

Query Match 38.7%; Score 70; DB 10; Length 31;

Best Local Similarity 100.0%; Pred. No. 0.00049;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITE 14



Db 18 QYIKANSKFIGITE 31

Search completed: December 4, 2002, 13:05:30

Job time : 2.93333 secs

Db 4 QYIKANSKFIGITE 17

|||||

RESULT 14

US-09-983-019-5

; Sequence 5, Application US/09983019

; Patent No. US20020146820A1

; GENERAL INFORMATION:

; APPLICANT: Diamond, Don J.

; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS PP15

; FILE REFERENCE: 1954-347US

; CURRENT APPLICATION NUMBER: US/09/983,019

; CURRENT FILING DATE: 2001-10-22

; PRIOR APPLICATION NUMBER: US 60/241,944

; PRIOR FILING DATE: 2000-10-20

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 5

; LENGTH: 31

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; NAME/KEY: misc.feature

; LOCATION: (1)..()

; OTHER INFORMATION: Human cytomegalovirus vaccine peptide

; NAME/KEY: LIPID

; LOCATION: (1)..(1)

; OTHER INFORMATION: palmitic acid

US-09-983-019-5

Query Match 38.7%; Score 70; DB 10; Length 31;

Best Local Similarity 100.0%; Pred. No. 0.00049;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITE 14

|||||

Db 4 QYIKANSKFIGITE 17

RESULT 15

US-09-983-019-6

; Sequence 6, Application US/09983019

; Patent No. US20020146820A1

; GENERAL INFORMATION:

; APPLICANT: Diamond, Don J.

; TITLE OF INVENTION: IMMUNOREACTIVE PEPTIDE CTL EPITOPES OF HUMAN CYTOMEGALOVIRUS PP15

; FILE REFERENCE: 1954-347US

; CURRENT APPLICATION NUMBER: US/09/983,019

; CURRENT FILING DATE: 2001-10-22

; PRIOR APPLICATION NUMBER: US 60/241,944

; PRIOR FILING DATE: 2000-10-20

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 6

; LENGTH: 31

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; NAME/KEY: misc.feature

; LOCATION: (1)..()

; OTHER INFORMATION: Human cytomegalovirus vaccine peptide

; NAME/KEY: LIPID

; LOCATION: (1)..(1)

; OTHER INFORMATION: palmitic acid

US-09-983-019-6

Query Match 38.7%; Score 70; DB 10; Length 31;

Best Local Similarity 100.0%; Pred. No. 0.00049;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYIKANSKFIGITE 14

|||||

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2002, 12:49:07 ; Search time 5.33333 Seconds  
(without alignments)  
176.538 Million cell updates/sec

Title: US-09-362-731A-5

Perfect score: 181

Sequence: 1 QYIKANSKFIGITELGGCHGSEPCNIHRGKPF 32

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length:	200000000
Maximum DB seq length:	200000000

.....

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

```
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
```

2: /cqn2\_6/ptodata/1/iaa/5B\_COMB.pcp.\*

3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pcp:\*

4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pcp:\*

5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.per

6: /cgn2\_6/ptodata/1/iaa/backfiles1.per

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query			DB	ID	Description
		Match	Length				
1	92.5	51.1	92	2	US-08-482-142-160	Sequence 160, App	
2	92.5	51.1	92	2	US-08-478-572-160	Sequence 160, App	
3	92.5	51.1	92	4	US-08-484-296-160	Sequence 42, Appl	
4	92	50.8	25	2	US-08-482-142-42	Sequence 42, Appl	
5	92	50.8	25	2	US-08-478-572-42	Sequence 42, Appl	
6	92	50.8	25	4	US-08-484-296-42	Sequence 42, Appl	
7	92	50.8	25	5	PCT-US95-04481-19	Sequence 19, Appl	
8	92	50.8	35	3	US-08-460-040-7	Sequence 7, Appl	
9	92	50.8	47	1	US-08-446-692-35	Sequence 35, Appl	
10	92	50.8	47	2	US-08-488-351A-35	Sequence 35, Appl	
11	92	50.8	61	2	US-08-482-142-162	Sequence 162, App	
12	92	50.8	61	2	US-08-478-572-162	Sequence 162, App	
13	92	50.8	61	4	US-08-484-296-162	Sequence 162, App	
14	92	50.8	84	2	US-08-482-142-161	Sequence 161, App	
15	92	50.8	84	2	US-08-478-572-161	Sequence 161, App	
16	92	50.8	84	4	US-08-484-296-161	Sequence 161, App	
17	92	50.8	129	1	US-07-945-288-12	Sequence 12, Appl	
18	92	50.8	129	1	US-08-462-831-12	Sequence 12, Appl	
19	92	50.8	129	1	US-08-461-809-12	Sequence 12, Appl	
20	92	50.8	129	1	US-08-461-441-12	Sequence 12, Appl	
21	92	50.8	129	2	US-08-482-142-157	Sequence 157, App	
22	92	50.8	129	2	US-08-482-142-159	Sequence 159, App	
23	92	50.8	129	2	US-08-478-572-157	Sequence 157, App	
24	92	50.8	129	2	US-08-478-572-159	Sequence 159, App	
25	92	50.8	129	4	US-08-484-296-157	Sequence 157, App	
26	92	50.8	129	4	US-08-484-296-159	Sequence 159, App	
27	92	50.8	129	5	PCT-US93-08518-12	Sequence 12, Appl	

28	92	50	8	2	US-08-482-142-158	Sequence 158, App
29	92	50	8	2	US-08-478-572-158	Sequence 158, App
30	92	50	8	2	US-08-484-296-158	Sequence 6, Appli
31	92	50	8	145	US-08-460-040-6	Sequence 4, Appli
32	92	50	8	146	US-07-945-288-4	Sequence 4, Appli
33	92	50	8	146	US-08-462-831-4	Sequence 4, Appli
34	92	50	8	146	US-08-461-809-4	Sequence 4, Appli
35	92	50	8	146	US-08-461-441-4	Sequence 4, Appli
36	92	50	8	146	US-08-482-142-4	Sequence 4, Appli
37	92	50	8	146	US-08-478-572-4	Sequence 4, Appli
38	92	50	8	146	US-08-484-296-4	Sequence 4, Appli
39	92	50	8	146	PCR-US93-08519-4	Sequence 4, Appli
40	90	49	7	129	US-08-330-264-18	Sequence 104, App
41	89	49	2	25	US-08-482-142-104	Sequence 104, App
42	89	49	2	25	US-08-478-572-104	Sequence 104, App
43	89	49	2	25	US-08-484-296-104	Sequence 8, Appli
44	89	49	2	129	US-07-945-288-8	Sequence 13, Appli
45	89	49	2	129	US-07-945-288-13	

## ALIGNMENTS

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match 51.1%; Score 92.5; DB 2; Length 92;  
 Best Local Similarity 67.9%; Pred. No. 8.3e-06;  
 Matches 19; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 5 ANSKFIGITELGGCHGSEPCNIHRGKPF 32  
 |||: | : ||||| |||||  
 Db 9 ANSE-IKKVMVPGCHGSEPCIHRGKPF 35

RESULT 2  
 US-08-478-572-160  
 ; Sequence 160, Application US/08478572  
 ; Patent No. 5968526  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Garman, Richard  
 ; APPLICANT: Greenstein, Julia  
 ; APPLICANT: Kuo, Mei-chang  
 ; APPLICANT: Rogers, Bruce  
 ; APPLICANT: Franzen, Henry  
 ; APPLICANT: Chen, Xian  
 ; APPLICANT: Evans, Sean  
 ; APPLICANT: Shaked, Ze'ev  
 ; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
 ; FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
 ; NUMBER OF SEQUENCES: 207  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
 ; STREET: 610 LINCOLN STREET  
 ; CITY: WALTHAM  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02154  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: ASCII TEXT  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/478,572  
 ; FILING DATE: 07-June-1995  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/445,307  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: CRAIG, ANNE I.  
 ; REGISTRATION NUMBER: 32,976  
 ; REFERENCE/DOCKET NUMBER: 017.605  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 466-6000  
 ; TELEFAX: (617) 466-6040  
 ; INFORMATION FOR SEQ ID NO: 160:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 92 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; FRAGMENT TYPE: internal  
 ; US-08-478-572-160

Query Match 51.1%; Score 92.5; DB 2; Length 92;  
 Best Local Similarity 67.9%; Pred. No. 8.3e-06;  
 Matches 19; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 5 ANSKFIGITELGGCHGSEPCNIHRGKPF 32  
 |||: | : ||||| |||||  
 Db 9 ANSE-IKKVMVPGCHGSEPCIHRGKPF 35

RESULT 3  
 US-08-484-296-160  
 ; Sequence 160, Application US/08484296

; Patent No. 6268491  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Garman, Richard  
 ; APPLICANT: Greenstein, Julia  
 ; APPLICANT: Kuo, Mei-chang  
 ; APPLICANT: Rogers, Bruce  
 ; APPLICANT: Franzen, Henry  
 ; APPLICANT: Chen, Xian  
 ; APPLICANT: Evans, Sean  
 ; APPLICANT: Shaked, Ze'ev  
 ; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
 ; FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
 ; NUMBER OF SEQUENCES: 207  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
 ; STREET: 610 LINCOLN STREET  
 ; CITY: WALTHAM  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02154  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: ASCII TEXT  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/484,296  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/445,307  
 ; FILING DATE: 07 June 1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: CRAIG, ANNE I.  
 ; REGISTRATION NUMBER: 32,976  
 ; REFERENCE/DOCKET NUMBER: 017.605  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 466-6000  
 ; TELEFAX: (617) 466-6040  
 ; INFORMATION FOR SEQ ID NO: 160:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 92 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; FRAGMENT TYPE: internal  
 ; US-08-484-296-160

Query Match 51.1%; Score 92.5; DB 4; Length 92;  
 Best Local Similarity 67.9%; Pred. No. 8.3e-06;  
 Matches 19; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 5 ANSKFIGITELGGCHGSEPCNIHRGKPF 32  
 |||: | : ||||| |||||  
 Db 9 ANSE-IKKVMVPGCHGSEPCIHRGKPF 35

RESULT 4  
 US-08-482-142-42  
 ; Sequence 42, Application US/08482142  
 ; Patent No. 5820862  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Garman, Richard  
 ; APPLICANT: Greenstein, Julia  
 ; APPLICANT: Kuo, Mei-chang  
 ; APPLICANT: Rogers, Bruce  
 ; APPLICANT: Franzen, Henry  
 ; APPLICANT: Chen, Xian  
 ; APPLICANT: Evans, Sean  
 ; APPLICANT: Shaked, Ze'ev  
 ; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
 ; FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
 ; NUMBER OF SEQUENCES: 207

```

CORRESPONDENCE ADDRESS:
ADDRESS: IMMUNOLOGIC PHARMACEUTICAL CORPORATION
STREET: 610 LINCOLN STREET
CITY: WALTHAM
STATE: MA
COUNTRY: USA
ZIP: 02154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII TEXT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/445,307
FILING DATE: 07 June 1995
ATTORNEY/AGENT INFORMATION:
NAME: CRAIG, ANNE I.
REGISTRATION NUMBER: 32,976
REFERENCE/DOCKET NUMBER: 017.6US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 466-6000
TELEFAX: (617) 466-6040
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
US-08-482-142-42

Query Match 50.8%; Score 92; DB 2; Length 25;
Best Local Similarity 93.8%; Pred. No. 2.6e-06;
Matches 15; Conservative 0; Mismatches 1; Indels

QY 17 GCHGSEPCNIHRGKPF 32
Db 10 GCHGSEPCNIHRGKPF 25

RESULT 5
US-08-478-572-42
Sequence 42, Application US/08478572
Patent No. 5968526
GENERAL INFORMATION:
APPLICANT: Garman, Richard
APPLICANT: Greenstein, Julia
APPLICANT: Kuo, Mei-chang
APPLICANT: Rogers, Bruce
APPLICANT: Franzen, Henry
APPLICANT: Chen, Xian
APPLICANT: Evans, Sean
APPLICANT: Shaked, Ze'ev
TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS
TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)
NUMBER OF SEQUENCES: 207
CORRESPONDENCE ADDRESS:
ADDRESS: IMMUNOLOGIC PHARMACEUTICAL CORPORATION
STREET: 610 LINCOLN STREET
CITY: WALTHAM
STATE: MA
COUNTRY: USA
ZIP: 02154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII TEXT
CURRENT APPLICATION DATA:

```

```
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
US-08-484-296-42

Query Match          50.8%; Score 92; DB 4; Length 25;
Best Local Similarity 93.8%; Pred. No. 2.6e-06;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 17 GCHGSEPCNIHRGKPF 32
      ||||| |||||
Db 10 GCHGSEPCIIHRGKPF 25

RESULT 7
PCT-US95-04481-19
; Sequence 19, Application PC/TUS9504481
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Pharmaceutical Peptide Formulations For Treatment of Dust Mit
; NUMBER OF SEQUENCES: 54
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04481
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/227,772
; FILING DATE: April 14, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Vanstone, Darlene A.
; REGISTRATION NUMBER: 35,279
; REFERENCE/DOCKET NUMBER: 017.5 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6040
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
PCT-US95-04481-19

Query Match          50.8%; Score 92; DB 5; Length 25;
Best Local Similarity 93.8%; Pred. No. 2.6e-06;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 17 GCHGSEPCNIHRGKPF 32
      ||||| |||||
Db 10 GCHGSEPCIIHRGKPF 25

RESULT 8
US-08-460-040-7
; Sequence 7, Application US/08460040
; Patent No. 6071522
; GENERAL INFORMATION:
; APPLICANT: Thomas, Wayne R.
; TITLE OF INVENTION: Cloning of Mite Allergens
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD

; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-460-040-7

Query Match          50.8%; Score 92; DB 3; Length 35;
Best Local Similarity 93.8%; Pred. No. 3.6e-06;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 17 GCHGSEPCNIHRGKPF 32
      ||||| |||||
Db 20 GCHGSEPCIIHRGKPF 35

RESULT 9
US-08-446-692-35
; Sequence 35, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 8-NOV-93
; APPLICATION NUMBER: 07/458,642
; FILING DATE: 13-FEB-90
; APPLICATION NUMBER: PCT/AU88/00195
; FILING DATE: 17-JUNE-88
; APPLICATION NUMBER: PI 2523/87
; FILING DATE: 18-JUNE-87
; ATTORNEY/AGENT INFORMATION:
; NAME: Amy E. Mandragouras
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: IMI-021CN2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-460-040-7
```

```
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Marla C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-446-692-35

Query Match 50.8%; Score 92; DB 1; Length 47;
Best Local Similarity 90.5%; Pred. No. 4.9e-06;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITELGGCHGS 21
Db 21 QYIKANSKFIGITELGGHWS 41

RESULT 10
US-08-488-351A-35
; Sequence 35, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Marla C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 35:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-488-351A-35

Query Match 50.8%; Score 92; DB 2; Length 47;
Best Local Similarity 90.5%; Pred. No. 4.9e-06;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITELGGCHGS 21
Db 21 QYIKANSKFIGITELGGHWS 41

RESULT 11
US-08-482-142-162
; Sequence 162, Application US/08482142
; Patent No. 5820862
; GENERAL INFORMATION:
; APPLICANT: Garman, Richard
; APPLICANT: Greenstein, Julia
; APPLICANT: Kuo, Mei-chang
; APPLICANT: Rogers, Bruce
; APPLICANT: Franzen, Henry
; APPLICANT: Chen, Xian
; APPLICANT: Evans, Sean
; APPLICANT: Shaked, Ze'ev
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS
; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,142
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,307
; FILING DATE: 07 June 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAIG, ANNE I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: 017,605
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6040
; INFORMATION FOR SEQ ID NO: 162:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 61 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-08-482-142-162

Query Match 50.8%; Score 92; DB 2; Length 61;
Best Local Similarity 93.8%; Pred. No. 6.4e-06;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 17 GCHGSEPCNIHRGKPF 32
Db 17 GCHGSEPCNIHRGKPF 32
```

Db 20 GCHGSEPCIIHRGKPF 35

## RESULT 12

US-08-478-572-162  
; Sequence 162, Application US/08478572  
; Patent No. 5968526

## ; GENERAL INFORMATION:

; APPLICANT: Garman, Richard  
; APPLICANT: Greenstein, Julia  
; APPLICANT: Kuo, Mei-chang  
; APPLICANT: Rogers, Bruce  
; APPLICANT: Franzen, Henry  
; APPLICANT: Chen, Xian  
; APPLICANT: Evans, Sean  
; APPLICANT: Shaked, Ze'ev  
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
; NUMBER OF SEQUENCES: 207  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
; STREET: 610 LINCOLN STREET  
; CITY: WALTHAM  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02154

## ; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/478,572  
; FILING DATE: 07-June-1995

## ; CLASSIFICATION:

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/445,307  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CRAIG, ANNE I.  
; REGISTRATION NUMBER: 32,976  
; REFERENCE/DOCKET NUMBER: 017.6US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 466-6000  
; TELEFAX: (617) 466-6040  
; INFORMATION FOR SEQ ID NO: 162:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 61 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; FRAGMENT TYPE: internal

US-08-478-572-162

Query Match 50.8%; Score 92; DB 2; Length 61;

Best Local Similarity 93.8%; Pred. No. 6.4e-06;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 17 GCHGSEPCIIHRGKPF 32

||||||| |||||

Db 20 GCHGSEPCIIHRGKPF 35

## RESULT 13

US-08-484-296-162  
; Sequence 162, Application US/08484296  
; Patent No. 6268491

## ; GENERAL INFORMATION:

; APPLICANT: Garman, Richard  
; APPLICANT: Greenstein, Julia  
; APPLICANT: Kuo, Mei-chang  
; APPLICANT: Rogers, Bruce  
; APPLICANT: Franzen, Henry

; APPLICANT: Chen, Xian  
; APPLICANT: Evans, Sean  
; APPLICANT: Shaked, Ze'ev  
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
; NUMBER OF SEQUENCES: 207  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
; STREET: 610 LINCOLN STREET  
; CITY: WALTHAM  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02154

## ; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,296  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/445,307  
; FILING DATE: 07 June 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CRAIG, ANNE I.  
; REGISTRATION NUMBER: 32,976  
; REFERENCE/DOCKET NUMBER: 017.6US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 466-6000  
; TELEFAX: (617) 466-6040  
; INFORMATION FOR SEQ ID NO: 162:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 61 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; FRAGMENT TYPE: internal

US-08-484-296-162

Query Match 50.8%; Score 92; DB 4; Length 61;

Best Local Similarity 93.8%; Pred. No. 6.4e-06;

Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 17 GCHGSEPCIIHRGKPF 32

||||||| |||||

Db 20 GCHGSEPCIIHRGKPF 35

## RESULT 14

US-08-482-142-161  
; Sequence 161, Application US/08482142  
; Patent No. 5820862

## ; GENERAL INFORMATION:

; APPLICANT: Garman, Richard  
; APPLICANT: Greenstein, Julia  
; APPLICANT: Kuo, Mei-chang  
; APPLICANT: Rogers, Bruce  
; APPLICANT: Franzen, Henry  
; APPLICANT: Chen, Xian  
; APPLICANT: Evans, Sean  
; APPLICANT: Shaked, Ze'ev  
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
; NUMBER OF SEQUENCES: 207  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
; STREET: 610 LINCOLN STREET  
; CITY: WALTHAM  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02154

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,142  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/445,307  
FILING DATE: 07 June 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: CRAIG, ANNE I.  
REGISTRATION NUMBER: 32,976  
REFERENCE/DOCKET NUMBER: 017.6US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 466-6000  
TELEFAX: (617) 466-6040  
INFORMATION FOR SEQ ID NO: 161:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 84 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-482-142-161

Query Match 50.8%; Score 92; DB 2; Length 84;  
Best Local Similarity 93.8%; Pred. No. 8.8e-06;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 17 GCHGSEPCNIHRGKPF 32  
||||||| |||||  
Db 20 GCHGSEPCIIHRGKPF 35

## RESULT 15

US-08-478-572-161  
Sequence 161, Application US/08478572  
Patent No. 5968526  
GENERAL INFORMATION:  
APPLICANT: Garman, Richard  
APPLICANT: Greenstein, Julia  
APPLICANT: Kuo, Mei-chang  
APPLICANT: Rogers, Bruce  
APPLICANT: Franzen, Henry  
APPLICANT: Chen, Xian  
APPLICANT: Evans, Sean  
APPLICANT: Shaked, Ze'ev  
TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
NUMBER OF SEQUENCES: 207  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
STREET: 610 LINCOLN STREET  
CITY: WALTHAM  
STATE: MA  
COUNTRY: USA  
ZIP: 02154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/478,572  
FILING DATE: 07-June-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/445,307  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:

NAME: CRAIG, ANNE I.  
REGISTRATION NUMBER: 32,976  
REFERENCE/DOCKET NUMBER: 017.6US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 466-6000  
TELEFAX: (617) 466-6040  
INFORMATION FOR SEQ ID NO: 161:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 84 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-478-572-161

Query Match 50.8%; Score 92; DB 2; Length 84;  
Best Local Similarity 93.8%; Pred. No. 8.8e-06;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 17 GCHGSEPCNIHRGKPF 32  
||||||| |||||  
Db 20 GCHGSEPCIIHRGKPF 35

Search completed: December 4, 2002, 12:55:09  
Job time : 5.33333 secs